



HAL
open science

Biochemical and structural exploration of the catalytic capacity of *Sulfolobus* KDG aldolases

Suzanne Wolterink-van Loo, André van Eerde, Marco A. J. Siemerink, Jasper Akerboom, Bauke W. Dijkstra, John van Der Oost

► **To cite this version:**

Suzanne Wolterink-van Loo, André van Eerde, Marco A. J. Siemerink, Jasper Akerboom, Bauke W. Dijkstra, et al.. Biochemical and structural exploration of the catalytic capacity of *Sulfolobus* KDG aldolases. *Biochemical Journal*, 2006, 403 (3), pp.421-430. 10.1042/BJ20061419 . hal-00478665

HAL Id: hal-00478665

<https://hal.science/hal-00478665v1>

Submitted on 30 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Biochemical and structural exploration of the catalytic capacity of *Sulfolobus* KDG aldolases

Suzanne Wolterink-van Loo*#, André van Eerde†#, Marco A. J. Siemerink*, Jasper Akerboom*,
Bauke W. Dijkstra†, and John van der Oost*

*Laboratory of Microbiology, Wageningen University, 6703 CT Wageningen, The Netherlands,
and †Laboratory of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG
Groningen, The Netherlands

Address correspondence to: Suzanne Wolterink–van Loo, 6703 CT Wageningen, The Netherlands
Tel: +31-317-483748; fax: +31-317-483829; Email: suzanne.vanloo@wur.nl

#Both authors contributed equally

Short page heading title: Exploration of *Sulfolobus* KDG aldolases

Keywords: KDG aldolase, thermophilic enzymes, *Sulfolobus*, carbon-carbon bond formation, substrate specificity

Abbreviations: KDG, 2-keto-3-deoxygluconate; KDGal, 2-keto-3-deoxygalactonate; KDPG, 2-keto-3-deoxy-6-phosphogluconate; KDGA, 2-keto-3-deoxygluconate aldolase; GA, glyceraldehyde; GAP, glyceraldehyde-3-phosphate; DHDPS, dihydrodipicolinate synthase; NAL, N-acetylneuraminic lyase; DHAP, dihydroxyacetone phosphate; FBP, fructose-bisphosphate; DERA, 2-deoxy-ribose-5-phosphate aldolase; rmsds, root-mean-square differences; KDO, 2-keto-3-deoxyoctonate.

ABSTRACT

Aldolases are enzymes with potential applications in biosynthesis, depending on their activity, specificity and stability. In this study, the genomes of *Sulfolobus* species were screened for aldolases. Two new 2-keto-3-deoxygluconate (KDG) aldolases from *S. acidocaldarius* and *S. tokodaii* were identified, overexpressed in *E. coli*, and characterised. Both enzymes were found to have biochemical properties similar to the previously characterized *S. solfataricus* KDG aldolase, including the condensation of pyruvate and either D,L-glyceraldehyde or D,L-glyceraldehyde-3-phosphate. The crystal structure of *S. acidocaldarius* KDG aldolase revealed the presence of a novel phosphate binding motif that allows the formation of multiple hydrogen bonding interactions with the acceptor substrate, and enables high activity with glyceraldehyde-3-phosphate. Activity analyses with unnatural substrates revealed that these three KDG aldolases readily accept aldehydes with 2-4 carbon atoms, and that even aldoses with 5 carbon atoms are accepted to some extent. Water-mediated interactions permit binding of substrates in multiple conformations in the spacious hydrophilic binding site, and correlate with the observed broad substrate specificity.

INTRODUCTION

Aldolases are among the few classes of enzymes capable of enlarging the carbon skeleton of molecules in a specific way by catalysing the condensation of a ketone donor and an aldehyde acceptor. These enzymes can be a tool for the specific production of carbon-carbon bonds in organic synthesis, since it is difficult to achieve stereochemical control with the widely used chemical aldol condensation reaction [1]. Aldolases have already successfully been used for the biosynthesis of a limited number of specialty compounds [2]. However, many of the currently used aldolases are expensive, unstable, limited in their catalytic range, and often require expensive substrates. The discovery of new aldolases or the engineering of known aldolases is therefore of continuing importance [2]. The availability of many sequenced genomes opens up new opportunities for identification of potentially suitable aldolases.

In nature, many different aldolases are encountered. Some aldolases play a vital physiological role in the degradation of metabolites, for instance during glycolysis; other aldolases are employed for carbon backbone assembly, e.g. in amino acid synthesis [3]. Two types of aldolase enzymes are distinguished based on their mechanisms. Type-I aldolases form a Schiff-base intermediate with the donor-substrate, whereas type-II aldolases are metal-dependent enzymes [1, 4]. Although both types can usually accept different acceptor substrates, they are rather specific for the donor substrate. Using a different classification, synthesizing aldolases are therefore also grouped according to their donor specificity: (1) dihydroxyacetone phosphate (DHAP)-dependent aldolases; (2) 2-deoxy-ribose-5-phosphate aldolase (DERA) – acetaldehyde dependent; (3) (phosphoenol)pyruvate-dependent aldolases; and (4) glycine-dependent aldolases [4]. Novel aldolases belonging to these classes may be found in organisms with divergent sugar metabolic pathways or degradation pathways of xenobiotics [2]. The thermostability of their proteins make extremophiles such as *Sulfolobus* particularly interesting sources of new enzymes.

Sulfolobus species are thermoacidophilic archaea that typically grow optimally at high temperatures (75-85 °C) and at a low pH (pH 2-4) [5-7]. Because of their global abundance [8] and ease of cultivation [6], they have become model archaea; genetic tools have been established [9-11] and the complete genome sequences of *S. solfataricus*, *S. tokodaii* and *S. acidocaldarius* have recently been unravelled [12-14]. In particular because of their resemblance to eukaryotic counterparts, several protein complexes from *Sulfolobus* have been selected for the analyses of fundamental processes, such as replication [15-17], transcription, and translation [18-20]. In addition, potential applications have been described for some of the thermostable *Sulfolobus* enzymes, including an unusual 2-keto-3-deoxygluconate (KDG) aldolase [21].

KDG aldolase from *Sulfolobus solfataricus* is a type-I aldolase and is a member of the N-acetylneuraminatase lyase (NAL) subfamily. This NAL subfamily consists of tetrameric enzymes that are specific for pyruvate as donor substrate but use different aldehydes as acceptor substrates. The *S. solfataricus* KDG aldolase was first characterised as an enzyme that catalyses the reversible conversion $\text{D,L-2-keto-3-deoxygluconate} \rightleftharpoons \text{pyruvate} + \text{D,L-glyceraldehyde (GA)}$. Later it was discovered that the enzyme was not only able to synthesize 2-keto-3-deoxygluconate but that it could also synthesize 2-keto-3-deoxygalactonate (KDGal) from pyruvate and D-glyceraldehyde [22]. Recently, Ahmed *et al.* (2005) showed that the enzyme also catalyses the interconversion $\text{D,L-2-keto-3-deoxy-6-phosphogluconate (KDPG)} \rightleftharpoons \text{pyruvate} + \text{D,L-glyceraldehyde-3-phosphate (GAP)}$ [23], suggesting that this bi-functional enzyme is part of a semi-phosphorylated Entner-Doudoroff pathway [23, 24]. Previously, the crystal structure of *S. solfataricus* KDGA (Sso-KDGA) was solved; in agreement with the above mentioned specificity studies, substrate soaking studies with KDG and KDGal revealed promiscuous binding with multiple conformations of these substrates due to different water-mediated interactions of the O-5 and O-6 hydroxyls in the rather spacious binding cavity [25].

In the present study, two hypothetical KDG aldolase genes from *S. acidocaldarius* and *S. tokodaii* were cloned, and after expression in *E. coli*, their biochemical properties were assessed, together with the previously characterised *S. solfataricus* KDG aldolase. In addition, the crystal structure of *S. acidocaldarius* KDG aldolase was determined, providing a molecular basis for the observed substrate specificity and stereoselectivity. Furthermore we have identified a novel phosphate binding motif in the KDG aldolases that explains their high activity with phosphorylated substrates.

MATERIALS AND METHODS

Cloning and expression

The hypothetical KDG aldolase genes from *S. acidocaldarius* (Saci_0225; Genbank Identifier 70606067) and *S. tokodaii* (ST2479; Genbank Identifier 15623602) were PCR-amplified from genomic DNA with primer pairs BG 1816 & BG 1817 and BG 1783 & 1784 (Table 1). The obtained products were restricted with NcoI and SalI and ligated into NcoI/SalI restricted pET24d (Novagen). These plasmids, designated pWUR193 (containing Saci_0225) and pWUR192 (containing ST2479) were transformed to *E. coli* BL21(DE3) and *E. coli* BL21(DE3)/pRIL. The clone pWUR122, containing the KDG-aldolase gene from *S. solfataricus* (SSO3197), described previously (Ahmed *et al.* 2005), was also transformed in *E. coli* BL21(DE3).

Purification

S. acidocaldarius KDG aldolase (Sac-KDGA) was produced by inoculating 2 liter LB (Luria-Bertani) medium supplemented with 50 µg/ml kanamycin with *E. coli* BL21(DE3) containing pWUR193. Subsequently, after 4 hours of growth at 37 °C, the culture was induced with 0.1 mM IPTG (isopropyl-β-D-thiogalactopyranoside) and growth was continued overnight. The cells were harvested by centrifugation (3800xg, 10 minutes), resuspended in 50 mM HEPES buffer pH 7, 20 mM KCl, and lysed using a French press. After centrifugation (27,000xg, 20 minutes) the soluble protein fraction was incubated at 75 °C for 30 minutes. This was centrifuged (48,000xg, 30 minutes) to obtain the heat stable cell free extract (HSCFE). The Sac-KDGA was further purified using anion-exchange chromatography (Q-sepharose; 20 mM Tris-HCl, pH 8.5, gradient 0-1 M NaCl), eluting at 0.30 M NaCl and gel filtration (Superdex™ 200 10/300 GL; 50 mM Tris/HCl, pH 7.8, 100 mM NaCl). *S. tokodaii* KDG aldolase (Sto-KDGA) was produced in the same way, except that the culture – inoculated with BL21(DE3)/pRIL containing pWUR192 – was induced with 0.02 mM IPTG on ice and further grown at 20 °C for 2 days (Q-sepharose-elution at 0.21 M NaCl). *S. solfataricus* KDG aldolase (Sso-KDGA) was also produced in a similar way as Sac-KDGA, but the cell pellet was resuspended in 90 mM HEPES, pH 7, 160 mM KCl and only purified by heat incubation and anion-exchange chromatography, eluting at 0.32 M NaCl. The enzyme purity was checked by SDS-PAGE, staining the proteins with Coomassie brilliant blue. Protein concentrations were determined with the BioRad – Bradford method.

Activity measurements

KDG aldolase activity was measured with the thiobarbituric acid (TBA) assay as described by Buchanan *et al.* [26]. In a standard aldolase assay the aldolase reaction was performed using the following conditions: 50 mM sodium phosphate buffer (pH 6), 50 mM sodium pyruvate, 20 mM D,L-glyceraldehyde and an appropriate amount of enzyme (~ 1 μg) were incubated at 70 °C for 10 minutes. The TBA-assay was then performed as described before [26]. V_{max} and K_m determinations were performed at optimal pH at 70 °C.

Aldolase activities with different aldehydes were measured by assaying the remaining pyruvate in a standardized lactate dehydrogenase (LDH) assay in triplo. Reactions of 100 μl volume containing 50 mM sodium phosphate buffer, pH 6, 10 mM pyruvate and 20 mM aldehyde and 1– 100 μg enzyme were incubated at 70 °C for 10-60 minutes. The reaction was stopped by adding 6 μl 20% TCA and precipitated protein was removed by centrifugation (16,000 g, 10 minutes). From 984 μl assay buffer (containing 100 mM Tris/HCl pH 7.5, 0.16 mM NADH and ~3 units LDH) the starting absorbance at 340 nm was read. After adding 16 μl sample and incubating for ~5 minutes, the absorbance was read again. The absorbance difference in NADH (molecular extinction coefficient 6.22 $\text{mM}^{-1} \text{cm}^{-1}$; [27]) was taken to be the pyruvate amount left. When aldose substrates were used, the TBA assay was also used.

Thermostability of the enzymes at 100°C was checked as follows. KDG aldolase was diluted to 40 $\mu\text{g}/\text{ml}$ in 50 mM sodium phosphate, pH 6, divided into 175 μl aliquots and closed into HPLC-vials, extra sealed with nylon. Vials were immersed in a 100 °C oil bath and sampled at different time points and immediately cooled on ice. Residual activity was measured with the standard TBA assay.

Crystallization

Sac-KDGA was crystallized using the hanging drop method. Hexagonal crystals of Sac-KDGA with typical dimensions of 0.3 x 0.3 x 0.2 mm were grown at 18°C. Drops were a mixture of 1 μl of a 19 mg/ml protein solution in water and 1 μl reservoir solution (0.1 M HEPES, pH 7.5, 30 % (v/v) PEG 400 and 0.2 M MgCl_2). Diffraction data of native crystals were collected at the ID14-1 beam line at the ESRF (Grenoble, France) at 100 K using the reservoir solution as a cryoprotectant. Data were processed with XDS [28] and programs from the CCP4 package [29], with different crystals being indexed either in space group $P3_121$ or in a $P6_522$ cell with related dimensions (Table 5). MOLREP [30] was first used to find a molecular replacement solution for the $P3_121$ data, using the coordinates of one Sso-KDGA subunit (PDB code 1w3i, [25]) as a search model. Afterwards, a molecular replacement solution was found in the second space group,

P6₅22, using the preliminary model of Sac-KDGA as a search model. The prime-and-switch protocol of RESOLVE [31] was used for density modification and removal of model bias, after which final models were obtained by iterative cycles of refinement in Refmac5 [32] and model building in Xtalview [33]. To obtain a Schiff-base intermediate with pyruvate, a crystal was transferred to a solution of 0.1 M MES buffer, pH 6.3, 32 % (v/v) PEG 400, 0.2 M MgCl₂ and 50 mM sodium pyruvate at 291 K and was flash-frozen after 20 minutes. Data were collected at the ID14-2 beamline at the ESRF (Grenoble, France) and processing was as before. Coordinates have been deposited with the Protein Data Bank for the native Sac-KDGA structure in *P3₁21* (PDB accession code 2NUW), in *P6₅22* (2NUX), and the complex with pyruvate (2NUY).

RESULTS & DISCUSSION

Sulfolobus genomes contain putative KDG aldolases

The three sequenced *Sulfolobus* genomes contain several open reading frames encoding hypothetical aldolases, *i.e.* enzymes that potentially catalyse aldol-cleavage or aldol-condensation reactions (Table 2). These include DHAP-dependent fructose-bisphosphate (FBP) aldolases [34], pyruvate-dependent dihydrodipicolinate synthases (DHDPS) [35] and KDG aldolases. DHDPS and KDG aldolases do not require (expensive) phosphorylated substrates, and are thus potentially interesting for industrial applications. Previous studies on the KDG aldolase from *S. solfataricus* suggested a broad substrate specificity [22, 24] and therefore we focussed on homologues of this enzyme. A BLAST analysis with the protein sequence of the *S. solfataricus* KDG aldolase (SSO3197) revealed high homology with the KDG aldolase from *Thermoproteus tenax* (Tte-KDGA) and hypothetical KDG aldolases from *S. acidocaldarius*, *S. tokodaii*, *Thermoplasma* and *Picrophilus torridus*. In particular, the active site residues (Lys-153, Tyr-129, Gly-177, Val-193, Pro-6) are conserved in these KDG aldolases, and also in some well-characterised enzymes of the NAL subfamily of type-I aldolases, such as the N-acetylneuraminate lyases from *Haemophilus influenzae* [36] and *E. coli* [37] and the DHDPS from *E. coli* [35] (Fig. 1). Based on this high similarity and a well-conserved gene context, in which the genes for the putative KDG aldolase, KDG kinase and GAPN-type dehydrogenase are clustered, the hypothetical aldolases from *S. acidocaldarius* and *S. tokodaii* are most likely part of a semi-phosphorylated Entner-Doudoroff pathway as was demonstrated for Sso-KDGA and Tte-KDGA [23]. Based on that, the here analyzed homologues were expected to have a similar activity with pyruvate and D,L-glyceraldehyde (phosphorylated and non-phosphorylated) as the KDG aldolase from *S. solfataricus* [23, 24, 26].

Stability and activity of Sulfolobus KDG aldolases

For a thorough characterisation of their properties the hypothetical KDG aldolase genes from *S. acidocaldarius* and *S. tokodaii* were cloned and expressed in *E. coli* using the pET expression system. The *S. solfataricus* KDG aldolase (Sso-KDGA) was produced using a previously generated pET24d-derived construct [23]. The three recombinant KDG aldolases were readily purified by heat incubation and subsequent anion-exchange chromatography. KDG aldolase activity was assayed in the synthesis direction with the substrates pyruvate and D,L-glyceraldehyde producing 2-keto-3-deoxygluconate (KDG) or its C4-epimer (2-keto-3-deoxygalactonate, KDGal). The recombinant enzymes from *S. acidocaldarius* and *S. tokodaii* had KDG aldolase activity in

the same range as the Sso-KDGA. For a thorough comparison between the three KDG aldolases, their pH and temperature optima, and Michaelis-Menten kinetics were determined. The biochemical properties of Sac-KDGA and Sto-KDGA are not very different from the previously studied Sso-KDGA (Table 3). All three enzymes had higher activity with GAP than with GA (Table 4), similar to reports by Ahmed *et al.* [23] and Lamble *et al.* [24] on Sso-KDGA activity. The stability of these enzymes at high temperature was also determined. The KDG aldolases show a high thermostability, although the half-life time of Sac-KDGA at 100°C is relatively short, which may be related to the optimal growth temperature of *S. acidocaldarius*, which is lower than that of the other two species [6]. Taken together, the biochemical experiments and the conserved gene context indicate that these enzymes are involved in a modified Entner-Doudoroff pathway, similar to what recently has been reported for the Sso-KDGA [23].

Crystal structure of Sac-KDGA

Structural information can provide detailed insights into the molecular basis of the catalytic properties of these aldolases. Sac-KDGA crystallized in two space groups, $P3_121$ and $P6_522$, for which diffraction data were obtained to a maximum resolution of 1.8 Å and 2.5 Å, respectively (Table 5). The structure could be determined by molecular replacement using the coordinates of previously elucidated Sso-KDGA ([25] 50% amino acid sequence identity) as the starting model. Initial electron density maps were of good quality and allowed the models in space groups $P3_121$ and $P6_522$ to be refined to acceptable *R* factors of 15.7 and 17.6 %, respectively (Table 5). Both space groups have two molecules in the asymmetric unit, with both of them containing all 288 residues of the native amino acid sequence. The total of four independently refined molecules are nearly identical with pair-wise root-mean-square differences (rmsds) of their C α atom positions of less than 0.20 Å. Sac-KDGA has a tetrameric structure (Fig. 2A) in which each subunit consists of a (β/α)₈ barrel with a C-terminal helical extension (Fig. 2B). The tetramer forms a ring-like structure with a hollow core of ~ 20 Å in diameter. The active site of each subunit is located at the C-terminal end of the β -barrel (Fig. 2B) and is only accessible from the inside of the tetrameric ring. This structure is not only similar to the structure of Sso-KDGA, as evidenced by the rmsd between 288 matching C α atoms of only 1.0 Å, but it is also similar to the structures of other crystallized members of the NAL subfamily, illustrating the structural conservation within this group of pyruvate-dependent aldolases [25].

The active site is among the most highly conserved regions in the protein. It consists of a pyruvate-binding cavity close to the catalytic Lys-153 (Fig. 3A), which is at the bottom of a wider

hydrophilic pocket, (Fig. 3B), accessible from within the tetrameric ring (Fig. 2B). The entrance to this hydrophilic pocket is somewhat constricted by the flexible side chains of two arginine residues, Arg-234 and Arg-105' from a neighbouring subunit (Fig. 2B, 3B). The Sac-KDGA active site residues adopt somewhat different conformations as a function of pH. Most conspicuously, at pH 7.5 the catalytic Lys-153 side chain displays some conformational variability. In its most populated conformation, its side chain points away from the pyruvate binding cavity and buries its amino-group into the interior of the protein in a hydrogen bond with the Ser-96 side chain (not shown). Since Sac-KDGA has its activity optimum around pH 6.5, data was also collected from a crystal that had been transferred to a cryosolution at pH 6.3 before freezing. At this pH, the lysine is in a conformation that would allow Schiff-base formation and catalysis (data not shown).

To assess the structure of the Schiff-base intermediate with pyruvate a crystal was soaked in a pyruvate-containing solution at pH 6.3, after which data was collected to 2.5 Å resolution. The well-defined pyruvate moiety is buried inside the active site and makes close interactions with backbone atoms of Thr-42, Thr-43, Val-193, Gly-177, and the ring of Pro-6, thereby ensuring the high specificity of the enzyme for pyruvate (Fig. 3A). This part of the active site and the pyruvate binding mechanism are highly conserved with Tyr-129 participating in both substrate binding and in the substrate-assisted catalysis, by relaying a proton between the pyruvate carboxylate and the aldehyde of the acceptor [25, 36].

Despite the different conformations of the Lys-153 side chain, the backbone conformations in the active site region of the Sac-KDGA structure are rigid and remain almost invariant upon pyruvate binding. As this rigidity was also observed for the Sso-KDGA structure [25], it may correlate with thermostability. Interestingly, the active sites of N-acetyl neuraminidase lyases from mesophilic *Haemophilus influenzae* and *E. coli* are nearly identical to the KDGA active sites [36, 37], indicating that the backbone conformation itself is not an adaptation to higher temperatures. The thermostable KDG aldolases appear to have stabilised this region, and therefore the protein, by rigidifying the loop that carries Tyr-129, which displays a higher flexibility in *H. influenzae* and *E. coli* NAL [36, 37].

In contrast to the strict requirement for pyruvate, the ability of KDGA to accept both 2-keto-3-deoxy(-6-phospho)gluconate as well as 2-keto-3-deoxy(-6-phospho)galactonate, in their D- as well as their L-enantiomer conformation, already indicates that the specificity for the acceptor substrate is much more relaxed than for the donor substrate. Previously, the structures of D-KDG and D-KDGal Schiff-base intermediates in Sso-KDGA [25] have revealed the binding mode for these non-phosphorylated substrates during catalysis. The binding modes of the carbon backbone

of the glyceraldehyde moieties show considerable variation across the independently determined structures and between D-KDG and D-KDGal. Still, in all cases the C-4 OH maintains a hydrogen bond with the Tyr-130 (Tyr-129 in Sac-KDGA) OH group (Fig. 3C), which is important for catalysis. The C-5 and C-6 OH groups show more flexibility and have different hydrogen bonding partners for D-KDG and D-KDGal [25]. These hydrogen bonding partners are conserved in the Sac-KDGA structure as well as in the sequence of Sto-KDGA (Fig. 1, 3), which is in agreement with their comparable aldolase activity on pyruvate and glyceraldehyde. No structural data is available for the binding of the L-enantiomers of both KDG and KDGal in the KDGA active site, but the hydrophilic nature of the pocket allows their flexible binding in a similar way as the D-enantiomers. This not only provides a structural explanation for the lack of stereocontrol of the KDG aldolases with their natural non-phosphorylated substrates, but also gives important clues about the applicability of these enzymes with non-physiological substrates.

Sulfolobus KDG aldolases contain a novel phosphate-binding motif

All KDG aldolases show higher activity with GAP than with GA (Table 4 and [23]), which may be of physiological relevance in the semi-phosphorylated Entner-Doudoroff metabolic pathway in *Sulfolobus* spp. It was suggested recently that starvation conditions may favour the slower non-phosphorylated route, whereas otherwise a KDG kinase may phosphorylate KDG to KDPG [24]. The possible binding modes of these larger phosphorylated substrates were investigated by using the Sso-KDGA and Sac-KDGA intermediate structures (Fig. 3C) as templates. In Sac-KDGA, the position of a bulky and highly polar phosphate group on the C-6 position is restricted on one side by the apolar side chains of Ile-239 (substituted by Leu in Sso- and Sto-KDGA) and Thr-42 (Fig. 3C). This side of the pocket would not be able to provide favourable interactions to the phosphate. Model building shows that minor rearrangements of the C-6 moiety would bring the phosphate closer to the entrance of the pocket and within hydrogen bonding distance of the polar side chains of Arg-234, Tyr-131 and in addition Arg-105' of a neighbouring subunit (Fig. 3D). These three residues are absolutely conserved among the KDG aldolases with confirmed KDPGA activity from *Sulfolobus* spp. and *T. tenax* (Fig. 1), and form a novel phosphate binding motif. In addition, water-mediated contacts would be possible with the nearby side chains of Ser-238, Asn-179 and Ser-195 (Fig. 3D). The promiscuity of binding of KDGA, which was already revealed for the glyceraldehyde moiety [25], is also present in the bound phosphate through these water-mediated contacts and the large conformational flexibility of the Arg-105' and Arg-234 side chains (Fig. 3D). Charge complementarity between the phosphate group and the two conserved arginine residues allows efficient targeting of GAP to the active site and additional enzyme-substrate

interactions may also increase substrate affinity. These additional interactions provide a structural explanation for the higher activity observed with GAP and the higher catalytic efficiency with KDPG than KDG. On the other hand, the phosphate group does not influence catalysis directly and does not significantly affect the conformation of the glyceraldehyde moiety (Fig. 3D), in line with the observation that the KDG aldolases do not absolutely require phosphorylated substrates.

The novel phosphate binding motif is only conserved among the KDG aldolases from *Sulfolobus* spp. and *T. tenax*. In putative KDG aldolases identified in other thermophilic archaea like *Thermoplasma* and *Picrophilus torridus*, the two arginine residues of the motif are not conserved (Fig. 1), neither is the genomic context similar to that of the *Sulfolobus* spp. Significantly, these organisms use a non-phosphorylative Entner-Doudoroff pathway [38, 39], in which their KDG aldolases do not require the dual KDGA/KDPGA functionality of KDG aldolases active in a semi-phosphorylative pathway. To our knowledge, the KDG aldolases from *Sulfolobus* spp. and *T. tenax* are the only members of the NAL subfamily identified so far that act on phosphorylated substrates and contain this motif. Therefore, the motif may be used to identify aldolases from this family with KDPGA activity.

Intriguingly, many of the enzymes of the $(\beta/\alpha)_8$ fold family that specifically act on phosphorylated substrates contain a conserved phosphate binding site [40], which differs entirely from the phosphate binding motif of the KDG aldolases. In these enzymes, phosphate moieties are often bound between the diverging ends of strands $\beta 7$ and $\beta 8$ by their backbone amide groups [40]. Our structural analysis reveals that this structural $\beta 7$ - $\beta 8$ feature is not conserved in the *Sulfolobus* KDG aldolases. In the Sac-KDGA structure, a depression exists between strands $\beta 7$ and $\beta 8$, where two water molecules reside, hydrogen bonding to Gly-177, the side chain of Ser-178 and the main chain amides and side chains of Asn-179 and Ser-195 (Fig. 3B). Manual docking indicates that a phosphate group cannot be accommodated at this position; the backbone amide groups in Sac-KDGA and Sso-KDGA are not oriented properly for an optimal interaction with the phosphoryl oxygen atoms and the side chain of Asn-179 in Sac-KDGA would give steric hindrance (Fig. 3B,D). In addition, Asn-179 is substituted by an aspartate residue in both Sso- and Sto-KDGA (Fig. 1), which would give unfavourable interactions with a negatively charged phosphate moiety. These structural differences lead us to conclude that this $\beta 7$ - $\beta 8$ site is not favourable for phosphate binding and support the novel phosphate binding mechanism that is outlined above.

The phosphorylated KDPG is also a substrate for the KDPG aldolases from *E. coli* and *Pseudomonas putida*. These type-I aldolases also have a $(\beta/\alpha)_8$ fold, but they do not belong to the

NAL subfamily and the details of their catalytic mechanisms are different [41, 42]. The KDPG aldolase from *E. coli* has e.g. a rmsd of 2.8 Å over 171 matching C α atoms with Sac-KDGA. Remarkably, these aldolases do contain a hydrophilic space between the β 7 and β 8 strands with the structural features of a phosphate binding site. Significantly, in both the *E. coli* and *P. putida* KDPG aldolase crystal structures a sulphate ion is bound at this β 7- β 8 site [41, 42]. The overall similarity between sulphate and phosphate ions strongly suggests that, in contrast to the *Sulfolobus* KDG aldolases, the β 7- β 8 site functions as a binding site for the phosphate group of their substrate. Interestingly, the KDPG aldolases from *E. coli* and *P. putida* do not perform well without the phosphate group on the glyceraldehyde moiety. This indicates that their active site confers considerably more spatial specificity towards their phosphorylated acceptor substrate than the open hydrophilic binding site of the *Sulfolobus* KDG aldolases. The strict specificity of the bacterial KDPG aldolases fits their biological function, but makes them less suitable for biotechnological applications. In contrast, the unique phosphate binding motif of the archaeal KDG aldolases incorporates the flexibility that is necessary for promiscuous substrate binding. This flexibility may extend to, or be easily adapted to, the conversion of a wider range of acceptor substrates.

Structural requirements for acceptor substrates of KDG aldolases

KDG aldolases display substrate promiscuity towards their natural aldehyde substrates [22-24]. It was reported recently that Sso-KDGA can also accept glyceraldehyde acetonide in the synthesis reaction [21]. Furthermore, the structures of Sac-KDGA and Sso-KDGA suggest that the substrate binding site is able to accept a wide range of acceptor substrates. To identify the structural requirements for aldehydes as acceptor, the activity of Sac-KDGA, Sso-KDGA and Sto-KDGA was assayed with a range of non-physiological aldehyde acceptors (Table 4, and results not shown). Like their other properties, the substrate specificities of the different *Sulfolobus* KDG aldolases are quite comparable and will not be discussed separately.

First, the importance of the C-6 moiety was investigated by testing the 2-carbon aldehydes glycolaldehyde and glyoxylate. The available structural information suggests that the binding site should be able to accommodate and orient these smaller substrates through hydrogen bonds with the C-5 hydroxyl group. Indeed, activity rates with these substrates were similar to or only slightly lower than that of glyceraldehyde (Table 4), establishing that the KDG aldolases do accept substrates with only two carbon atoms. The reduced activity with glyoxylate may be the result of its resemblance to pyruvate and weak competitive inhibition of Schiff-base formation. In

comparison, KDPG aldolases from *P. putida*, *E. coli* and *Zymomonas mobilis* [43] react quite well with glyoxylate, but react with very low rates only with glycolaldehyde. This provides additional evidence that the KDG aldolases from *Sulfolobus* can perform well without a negative charge on the substrate and establishes that they have a catalytic range which is distinctly different from the bacterial KDPG aldolases.

After having established that the C-6 hydroxyl group is dispensable, the importance of the hydroxyl group at the C-5 position was investigated. In the Sso-KDGA KDG and KDGal complexes, this hydroxyl group has hydrogen bonding interactions with either Tyr-131 or water-mediated contacts with Gly-177 and Ser-195 ([25], see also Fig. 3C). Two aldehydes were tested that do not contain such a hydroxyl group at the C-5, and this showed that acetaldehyde is not accepted and propionaldehyde only at a very low rate (<0.1% of glyceraldehyde) (Table 4). This suggests that the polar hydroxyl substituent at the C-5 position is indispensable for proper binding and orientation of the aldehyde in the open hydrophilic binding pocket.

Next, the ability of the KDG aldolases to accept aldehydes with one more carbon atom than their natural substrates was explored using the simple four-carbon aldotetroses in all possible stereoisomers. The large hydrophilic binding pocket should be able to provide favourable interactions for the additional hydroxyl groups of these sugars. Indeed, all aldotetroses tested were readily accepted in synthesis reactions at 5-25 % of the glyceraldehyde rate. As all stereoisomers are accepted, albeit with a substantially higher rate for the reaction with L-threose, the KDG aldolases show a limited enantioselectivity on these larger substrates. The relative rates are again higher than those of the different KDPG aldolases of *P. putida*, *E. coli* and *Z. mobilis* [43], which accepted these compounds at 0-0.5% of the GAP rate only.

In contrast, the larger aldopentoses that were tested were accepted at low rates (<0.5% of GA) (Table 4). The binding site should however be large enough to accommodate these five-carbon acceptors. The low rates with the aldopentoses are likely due to a reduced effective concentration of the reactive open-chain conformation of these sugars, which is only ~ 0.1 % of the total aldopentose concentration at 28°C [44]. Significantly, the activity of the cleavage of the 2-keto-3-deoxyoctonate (KDO) (5 mM) to D-arabinose and pyruvate, is much higher than its condensation, reaching 0.146 ±0.006 U/mg (*S. acidocaldarius*), 0.470 ±0.040 U/mg (*S. solfataricus*) and 0.393 ±0.025 U/mg (*S. tokodaii*). Activity appears to be correlated with the effective concentration of the open-chain forms of the substrates, which is probably much higher for the ketone KDO than for the aldehyde D-arabinose, suggesting that the KDG aldolases do not catalyse the ring opening step of sugar substrates. Ring opening as the rate-limiting step may also

play a role with the aldotetroses [45], and product cyclization may effectively limit the rate of the reverse aldol reaction [24, 46].

Taken together, these results define the substrate requirements, or lack thereof, of these KDG aldolases. The substrate requirements are primarily restricted to the substituents of the carbon atom immediately adjacent to the aldehyde moiety, making glycolaldehyde the smallest acceptor substrate that can still be used by the KDG aldolases. Despite the absence of enantioselectivity in the condensation reaction, the hydroxyl group at the C-5 position of the synthesis product is indispensable for catalysis by providing a significant binding affinity of the aldehyde substrate. Our measurements do not rule out the possibility that additional hydrogen bonding interactions of other parts of an aldehyde acceptor may partially compensate for the decrease in binding affinity. The structure of the binding sites of Sac-KDGA and Sso-KDGA combined with our activity experiments explain the general lack of specific requirements on the structure of the acceptor beyond the glycolaldehyde moiety. This suggests that these enzymes should be applicable in the biosynthesis of a broad range of compounds which carry this structural motif.

The three different *Sulfolobus* KDG aldolases are not the only NAL family members which accept several different (non-phosphorylated) aldehydes in the condensation reaction with pyruvate. Other members of the NAL subfamily also have broad acceptor specificity, indicating the general versatility of their binding sites. Strikingly, although their binding sites are very similar, acceptors are bound very differently in different enzymes. In N-acetylneuraminase lyase e.g., substrate analogues may be accommodated in a different part of the active site pocket [36]. Given the similarity between active sites and substrates it was previously suggested that the substrates sialic acid and KDG would bind in a similar way and that substrate specificity within the NAL subfamily would be modulated in a subtle way by only a few amino acid mutations [36]. The structural and biochemical evidence on Sac-KDGA and Sso-KDGA [25] now give striking illustrations of how structurally similar binding sites may bind substrates in entirely different ways. This suggests that a limited number of mutations in the active site may already lead to novel substrate specificities. This has recently been demonstrated by the creation of NAL enzymes with DHDPS activity [47] or with L-KDO aldolase activity [48]. The thermostability, ease of production and well-characterised acceptor specificity of the *Sulfolobus* KDG aldolases make these enzymes highly suitable platforms for such an approach.

ACKNOWLEDGEMENTS

This research is performed as part of the IBOS Programme (Integration of Biosynthesis & Organic Synthesis) of Advanced Chemical Technologies for Sustainability (ACTS). In addition we want to thank Theo Sonke (DSM, Geleen) and Maurice Franssen (Wageningen University) for fruitful discussions.

REFERENCES

- 1 Machajewski, T. D. and Wong, C. H. (2000) The catalytic asymmetric aldol reaction. *Angew. Chem. Int. Ed. Engl.* **39**, 1352-1375
- 2 Samland, A. K. and Sprenger, G. A. (2006) Microbial aldolases as C-C bonding enzymes-- unknown treasures and new developments. *Appl. Microbiol. Biotechnol.* **71**, 253-264
- 3 Gijzen, H. J., Qiao, L., Fitz, W. and Wong, C. H. (1996) Recent advances in the chemoenzymatic synthesis of carbohydrates and carbohydrate mimetics. *Chem. Rev.* **96**, 443-474
- 4 Takayama, S., McGarvey, G. J. and Wong, C. H. (1997) Microbial aldolases and transketolases: new biocatalytic approaches to simple and complex sugars. *Annu. Rev. Microbiol.* **51**, 285-310
- 5 Brock, T. D., Brock, K. M., Belly, R. T. and Weiss, R. L. (1972) *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch. Mikrobiol.* **84**, 54-68
- 6 Grogan, D. W. (1989) Phenotypic characterization of the archaeobacterial genus *Sulfolobus*: comparison of five wild-type strains. *J. Bacteriol.* **171**, 6710-6719
- 7 Suzuki, T., Iwasaki, T., Uzawa, T., Hara, K., Nemoto, N., Kon, T., Ueki, T., Yamagishi, A. and Oshima, T. (2002) *Sulfolobus tokodaii* sp. nov. (f. *Sulfolobus* sp. strain 7), a new member of the genus *Sulfolobus* isolated from Beppu Hot Springs, Japan. *Extremophiles* **6**, 39-44
- 8 DeLong, E. F. and Pace, N. R. (2001) Environmental diversity of bacteria and archaea. *Syst. Biol.* **50**, 470-478
- 9 Stedman, K. M., Schleper, C., Rumpf, E. and Zillig, W. (1999) Genetic requirements for the function of the archaeal virus SSV1 in *Sulfolobus solfataricus*: construction and testing of viral shuttle vectors. *Genetics* **152**, 1397-1405
- 10 Aravalli, R. N. and Garrett, R. A. (1997) Shuttle vectors for hyperthermophilic archaea. *Extremophiles* **1**, 183-191
- 11 Cannio, R., Contursi, P., Rossi, M. and Bartolucci, S. (1998) An autonomously replicating transforming vector for *Sulfolobus solfataricus*. *J. Bacteriol.* **180**, 3237-3240
- 12 She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I. G., Curtis, B. A., De Moors, A., *et al.* (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 7835-7840
- 13 Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankaei, A., Kosugi, H., Hosoyama, *et al.* (2001) Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain 7. *DNA Res.* **8**, 123-140
- 14 Chen, L., Brugger, K., Skovgaard, M., Redder, P., She, Q., Torarinsson, E., Greve, B., Awayez, M., Zibat, A., Klenk, H. P. *et al.* (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. *J. Bacteriol.* **187**, 4992-4999
- 15 Klimczak, L. J., Grummt, F. and Burger, K. J. (1985) Purification and characterization of DNA polymerase from the archaeobacterium *Sulfolobus acidocaldarius*. *Nucleic Acids Res.* **13**, 5269-5282
- 16 Rossi, M., Rella, R., Pensa, M., Bartolucci, S., De Rosa, M., Gambacorta, A., Raia, C. A. and Dell'Aversano Orabona, N. (1986) Structure and properties of a thermophilic and thermostable DNA polymerase isolated from *Sulfolobus solfataricus*. *Syst. Appl. Microbiol.* **7**, 337-341
- 17 Dionne, I., Robinson, N. P., McGeoch, A. T., Marsh, V. L., Reddish, A. and Bell, S. D. (2003) DNA replication in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Biochem. Soc. Trans.* **31**, 674-676
- 18 Zillig, W., Stetter, K. O. and Janekovic, D. (1979) DNA-dependent RNA polymerase from the archaeobacterium *Sulfolobus acidocaldarius*. *Eur. J. Biochem.* **96**, 597-604
- 19 Reiter, W. D., Palm, P. and Zillig, W. (1988) Analysis of transcription in the archaeobacterium *Sulfolobus* indicates that archaeobacterial promoters are homologous to eukaryotic pol II promoters. *Nucleic Acids Res.* **16**, 1-19
- 20 Bell, S. D. and Jackson, S. P. (1998) Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. *Trends Microbiol.* **6**, 222-228
- 21 Lamble, H. J., Danson, M. J., Hough, D. W. and Bull, S. D. (2005) Engineering stereocontrol into an aldolase-catalysed reaction. *Chem. Commun. (Camb.)*, 124-126

- 22 Lambie, H. J., Heyer, N. I., Bull, S. D., Hough, D. W. and Danson, M. J. (2003) Metabolic pathway promiscuity in the archaeon *Sulfolobus solfataricus* revealed by studies on glucose dehydrogenase and 2-keto-3-deoxygluconate aldolase. *J. Biol. Chem.* **278**, 34066-34072
- 23 Ahmed, H., Ettema, T. J., Tjaden, B., Geerling, A. C., van der Oost, J. and Siebers, B. (2005) The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic archaea: a re-evaluation. *Biochem. J.* **390**, 529-540
- 24 Lambie, H. J., Theodossis, A., Milburn, C. C., Taylor, G. L., Bull, S. D., Hough, D. W. and Danson, M. J. (2005) Promiscuity in the part-phosphorylative Entner-Doudoroff pathway of the archaeon *Sulfolobus solfataricus*. *FEBS Lett.* **579**, 6865-6869
- 25 Theodossis, A., Walden, H., Westwick, E. J., Connaris, H., Lambie, H. J., Hough, D. W., Danson, M. J. and Taylor, G. L. (2004) The structural basis for substrate promiscuity in 2-keto-3-deoxygluconate aldolase from the Entner-Doudoroff pathway in *Sulfolobus solfataricus*. *J. Biol. Chem.* **279**, 43886-43892
- 26 Buchanan, C. L., Connaris, H., Danson, M. J., Reeve, C. D. and Hough, D. W. (1999) An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. *Biochem. J.* **343**, 563-570
- 27 Horecker, B. L. and Kornberg, A. (1948) The extinction coefficients of the reduced band of pyridin nucleotides. *J. Biol. Chem.* **175**, 385-390
- 28 Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795-800
- 29 Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760-763
- 30 Vagin, A. and Teplyakov, A. (1997) MOLREP: An automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022-1025
- 31 Terwilliger, T. C. (2004) Using prime-and-switch phasing to reduce model bias in molecular replacement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2144-2149
- 32 Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. and Dodson, E. J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 247-255
- 33 McRae, D. E. (1999) XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156-165
- 34 Siebers, B., Brinkmann, H., Dorr, C., Tjaden, B., Lilie, H., van der Oost, J. and Verhees, C. H. (2001) Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase. *J. Biol. Chem.* **276**, 28710-28718
- 35 Dobson, R. C., Valegard, K. and Gerrard, J. A. (2004) The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: further evidence for a catalytic triad. *J. Mol. Biol.* **338**, 329-339
- 36 Barbosa, J. A., Smith, B. J., DeGori, R., Ooi, H. C., Marcuccio, S. M., Campi, E. M., Jackson, W. R., Brossmer, R., Sommer, M. and Lawrence, M. C. (2000) Active site modulation in the N-acetylneuraminase lyase sub-family as revealed by the structure of the inhibitor-complexed *Haemophilus influenzae* enzyme. *J. Mol. Biol.* **303**, 405-421
- 37 Izard, T., Lawrence, M. C., Malby, R. L., Lilley, G. G. and Colman, P. M. (1994) The three-dimensional structure of N-acetylneuraminase lyase from *Escherichia coli*. *Structure* **2**, 361-369
- 38 Reher, M. and Schonheit, P. (2006) Glyceraldehyde dehydrogenases from the thermoacidophilic euryarchaeota *Picrophilus torridus* and *Thermoplasma acidophilum*, key enzymes of the non-phosphorylative Entner-Doudoroff pathway, constitute a novel enzyme family within the aldehyde dehydrogenase superfamily. *FEBS Lett.* **580**, 1198-1204
- 39 Budgen, N. and Danson, M. J. (1986) Metabolism of sugar via a modified Entner-Doudoroff pathway in the thermoacidophilic archaebacterium *Thermoplasma acidophilum*. *FEBS Lett.* **196**, 207-210
- 40 Nagano, N., Orengo, C. A. and Thornton, J. M. (2002) One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J. Mol. Biol.* **321**, 741-765

- 41 Allard, J., Grochulski, P. and Sygusch, J. (2001) Covalent intermediate trapped in 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase structure at 1.95-Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3679-3684
- 42 Bell, B. J., Watanabe, L., Rios-Steiner, J. L., Tulinsky, A., Lebioda, L. and Arni, R. K. (2003) Structure of 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase from *Pseudomonas putida*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **59**, 1454-1458
- 43 Shelton, M. C., Cotterill, I. C., Novak, S. T. A., Poonawala, R. M., Sudarshan, S. and Toone, E. J. (1996) 2-Keto-3-deoxy-6-phosphogluconate aldolases as catalysts for stereocontrolled carbon-carbon bond formation. *J. Am. Chem. Soc.* **118**, 2117-2125
- 44 Drew, K. N., Zajicek, J., Bondo, G., Bose, B. and Serianni, A. S. (1998) ¹³C-labeled aldopentoses: detection and quantitation of cyclic and acyclic forms by heteronuclear 1D and 2D NMR spectroscopy. *Carbohydr. Res.* **307**, 199-209
- 45 Serianni, A. S., Pierce, J., Huang, S.-G. and Barker, R. (1982) Anomerization of furanose sugars: kinetics of ring-opening reactions by ¹H and ¹³C saturation-transfer NMR spectroscopy. *J. Am. Chem. Soc.* **104**, 4037-4044
- 46 Midelfort, C. F., Gupta, R. K. and Meloche, H. P. (1977) Specificity of 2-keto-3-deoxygluconate-6-P aldolase for open chain form of 2-keto-3-deoxygluconate-6-P. *J. Biol. Chem.* **252**, 3486-3492
- 47 Joerger, A. C., Mayer, S. and Fersht, A. R. (2003) Mimicking natural evolution in vitro: an N-acetylneuraminidase mutant with an increased dihydrodipicolinate synthase activity. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 5694-5699
- 48 Hsu, C. C., Hong, Z., Wada, M., Franke, D. and Wong, C. H. (2005) Directed evolution of D-sialic acid aldolase to L-3-deoxy-manno-2-octulosonic acid (L-KDO) aldolase. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 9122-9126

Table 1.

Primers that are used in this study

Primer	Sequence
BG1784	CGCGCCCATGGTGTTCAAAATTTTAAGTATGGATATTG
BG1816	CGCGCGTCGACTTTACGAAACAGCTCTTTCTATTTTC
BG1817	CGCGCCATGGAAATAATTTACCTATCATTAC
BG1852	GCGCGTCGACTTAATGTACCAGTTCTTGAATCTTTC

Table 2.

Overview of possible enzymes -in *Sulfolobus* species- hypothetically catalyzing aldolase reactions

Enzyme	Function	Reaction
DAHP aldolase	Synthesis	PEP + Ery-4-P → 7P-2-dehydro-3-deoxy-D-arabinoheptonate
Citrate lyase	Degradation	Citrate → Oxaloacetate + acetate
DHDPS-like*	Synthesis	Pyruvate + Aspartate-β-semialdehyde → DHDP
FBP-aldolase ¹	Degradation	D-fructose 1,6-bisphosphate → dihydroxyacetone phosphate (DHAP) + D-glyceraldehyde 3- phosphate.
Isocitrate lyase ²	Degradation	isocitrate → succinate + glyoxylate
KDGA	Degradation	KDG → pyruvate + glyceraldehyde

* 3 paralogs in *S. solfataricus* 1 in *S. tokodaii* and *S. acidocaldarius*

¹ In *S. solfataricus* and *S. tokodaii* [34]

² In *S. solfataricus* and *S. acidocaldarius*

Table 3.

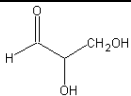
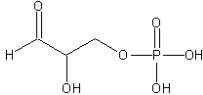



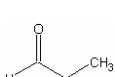
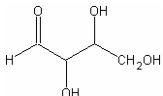
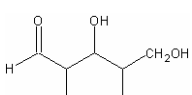
Comparison of biochemical data of *Sulfolobus* KDG aldolases

	<i>S. acidocaldarius</i>	<i>S. solfataricus</i>	<i>S. tokodaii</i>
Opt pH	6.5	6*	5.5
Opt. Temp	~99°C	95°C	90°C
$V_{max(\text{pyr})}$ (U/mg)	26.3 (\pm 0.9)	15.7 (\pm 0.3)*	17.8 (\pm 0.7)
$K_m(\text{pyr})$ (mM)	1.1 (\pm 0.2)	1.0 (\pm 0.1)*	0.7 (\pm 0.1)
$V_{max(\text{GA})}$ (U/mg)	33.1 (\pm 1.3)	17.1 (\pm 0.4)*	20.3 (\pm 0.8)
$K_m(\text{GA})$ (mM)	6.3 (\pm 0.5)	5.2 (\pm 0.1)*	3.9 (\pm 0.4)
$T_{1/2}$ (minutes at 100°C)	18.9 (\pm 2.1)	120 (150*)	110

* Data from Buchanan et al. [26]

Table 4.

 Substrate specificity of *Sulfolobus* KDGA aldolases based on the rate of condensation with pyruvate as donor substrate (U/mg), standard deviations are given in parenthesis

Aldehyde acceptor		Sac-KDGA*		Sso-KDGA*	
(D,L)-glyceraldehyde		37.3	(2.0)	15.1	(0.7)
(D,L)-glyceraldehyde-3-P		93.0	(4.6)	46.5	(0.9)
glycolaldehyde		40.6	(1.0)	10.1	(0.7)
glyoxylate		16.4	(0.9)	11.6	(0.2)
acetaldehyde		0		0	
propionaldehyde		0.0055	(0.0007)	0.0058	(0.0011)
(D)-erythrose		5.7	(0.2)	1.4	(0.1)
(L)-erythrose		2.3	(0.1)	2.0	(0.2)
(D)-threose		2.5	(0.1)	1.9	(0.1)
(L)-threose		10.6	(0.5)	2.3	(0.2)
(D)-ribose		0.028	(0.003)	0.015	(0.002)
(L)-arabinose		0.025	(0.002)	0.020	(0.003)
(D)-arabinose		0.010	(0.003)	0.023	(0.002)
(D)-xylose		0.050	(0.003)	0.034	(0.005)

* similar results were obtained with Sto-KDGA

Table 5.

Crystallographic data collection and refinement statistics on Sac-KDGA

	Native, pH 7.5	Native, pH 7.5	Pyruvate, pH 6.3
Space group	<i>P</i> 3 ₁ 21	<i>P</i> 6 ₅ 22	<i>P</i> 3 ₁ 21
Unit cell dimensions (Å)	<i>a</i> = <i>b</i> =108.2 <i>c</i> =171.9	<i>a</i> = <i>b</i> =109.5 <i>c</i> = 319.6	<i>a</i> = <i>b</i> =109.1 <i>c</i> =171.3
Resolution (Å)	30-1.8(1.88-1.8)	36-2.5(2.59-2.5)	50-2.5(2.65-2.5)
Wavelength (Å)	0.934	0.934	0.933
Unique reflections	107995	36273	39118
Completeness (%)	99.9 (99.8)	90.1 (55.2)	99.6 (97.1)
Redundancy	10.9 (4.0)	33.1(19.8)	6.2 (4.2)
<i>I</i> / σ (<i>I</i>)	14.1 (2.3)	23.8 (8.4)	17.6 (5.3)
<i>R</i> _{sym} ^a (%)	9.6 (55.9) ^c	15.5 (37.8)	5.3 (34.8)
Refinement			
<i>R</i> _{cryst} ^b (%)	15.7	17.6	15.9
<i>R</i> _{free} ^b (%)	17.9	21.7	19.1
# atoms protein	4634	4598	4616
pyruvate	-	-	10
solvent	412	95	162
Avg. <i>B</i> factor (Å ²) protein	35.5	52.0	52.9
pyruvate	-	-	42.6
solvent	44.2	43.6	50.8
R.m.s. deviations bonds/angles (Å/°)	0.012/1.5	0.008/1.3	0.008/1.3

Values in parentheses are for the highest resolution shell

^a*R*_{sym} = $|I - \langle I \rangle| / I$, where *I* is the observed intensity and $\langle I \rangle$ the average intensity

^b*R*_{cryst} = $\frac{\sum_{hkl, work} ||F_{obs}| - k|F_{calc}||}{\sum_{hkl} |F_{obs}|} * 100\%$, where *F*_{obs} = observed structure factor and *F*_{calc} = calculated structure factor. *R*_{free} = *R* calculated with 5% of randomly selected data that were omitted from the refinement.

^cValue for the highest resolution shell is affected by anisotropy in diffraction

FIGURE LEGENDS

Figure 1.

Alignment of the three *Sulfolobus* KDG aldolases with *Thermoproteus tenax* KD(P)G aldolase, hypothetical KDG aldolases from *Thermoplasma acidophilum*, *T. volcanium* and *Picrophilus torridus*, N-acetyl neuraminidase lyases from *E. coli* and *Haemophilus influenzae* and the *E. coli* dihydrodipicolinate synthase. Residues that are conserved between the KD(P)G aldolases are indicated with dark shading, while those residues that are also conserved in the hypotheticals and other aldolases are shaded lighter. At the arrow, the conserved catalytic lysine residue. Catalytic residues (▼) and residues forming the putative phosphate binding motif in Sac-KDGA are indicated (◆).

Figure 2.

Overall structure of Sac-KDGA (A) Sac-KDGA tetramer (B) Orthogonal views of the Sac-KDGA monomer. The Schiff-base forming Lys-153 is indicated in stick presentation with pyruvate (yellow) attached. To show the layout of the substrate binding pocket, the loop of the neighbouring subunit carrying Arg-105 is also shown, together with Arg-105 and Arg-234 (orange). (C) General scheme of the aldol condensation reaction catalyzed by the KDG aldolases.

Figure 3.

Stereo views of the Sac-KDGA active site (A) Weighted difference density map of pyruvate (green) bound to Lys-153 in the pyruvate-binding cavity of Sac-KDGA. Map is contoured at 4σ . (B) The substrate binding pocket of Sac-KDGA with pyruvate. Strands $\beta 7$ and $\beta 8$ are indicated. (C) Superposition of the Pyr:Sac-KDGA (green:grey) and KDG:Sso-KDGA complex ([25], yellow:cyan). Hydrogen bonding interactions with the pyruvate moiety have been omitted for clarity and the orientation is slightly different from B. Water molecules shown are from the KDG:Sso-KDGA complex. (D) KDPG (yellow) modelled in the active site of Sac-KDGA on the basis of the KDG:Sso-KDGA complex [25]. Water molecules from the crystal structure of native Sac-KDGA, which would have favourable interactions with the ligand are also shown. Some hydrogen bonding interactions have been omitted for clarity and the orientation is slightly different from B and C.

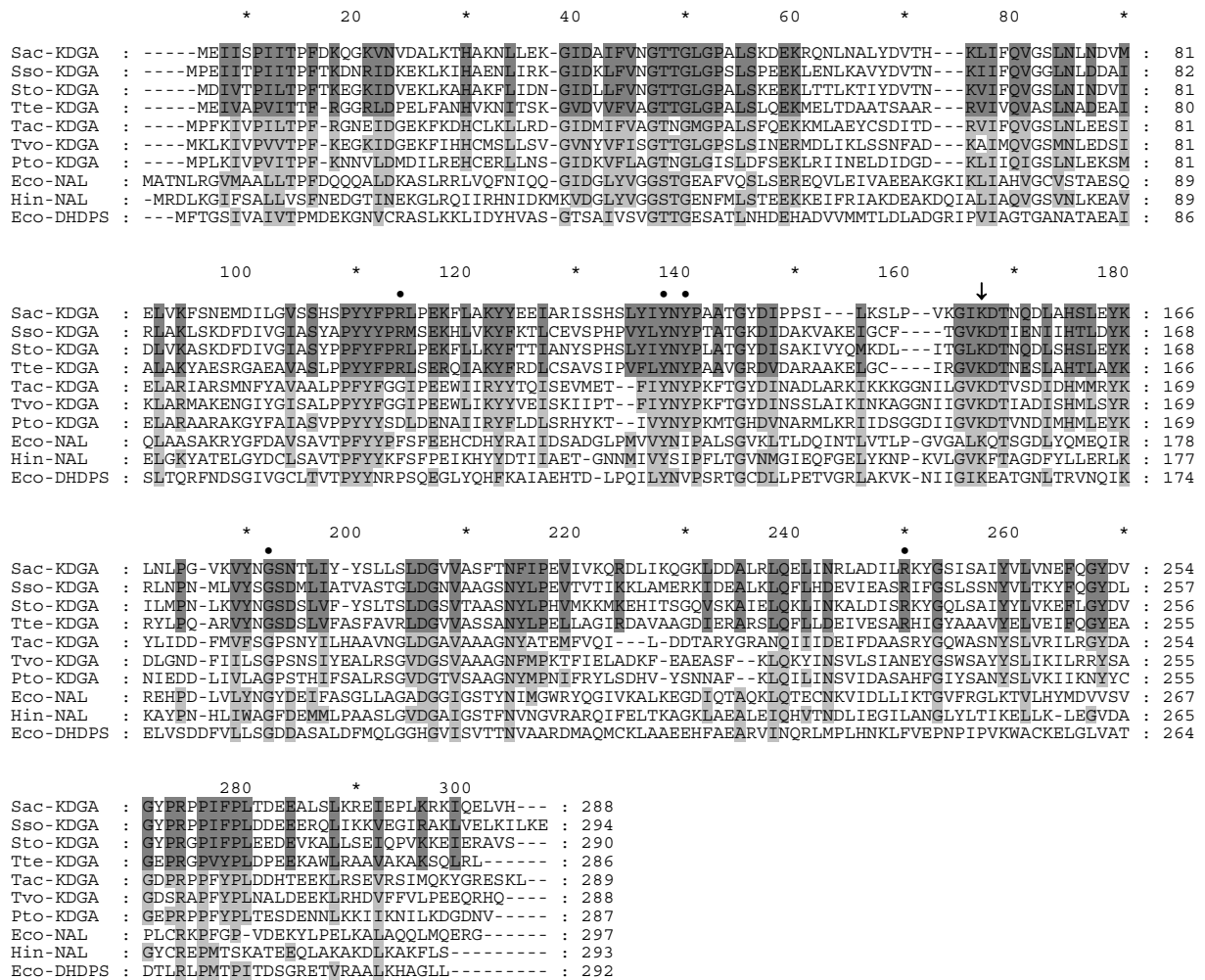


Figure 1.

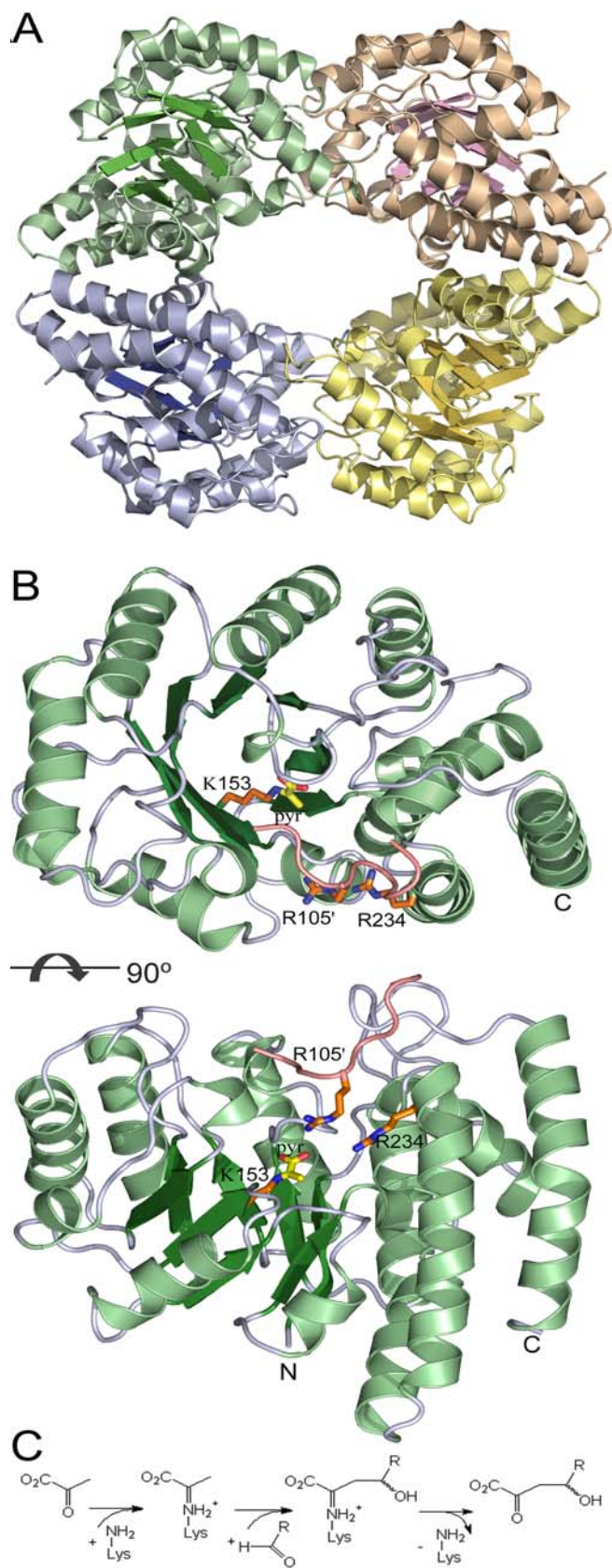


Figure 2.

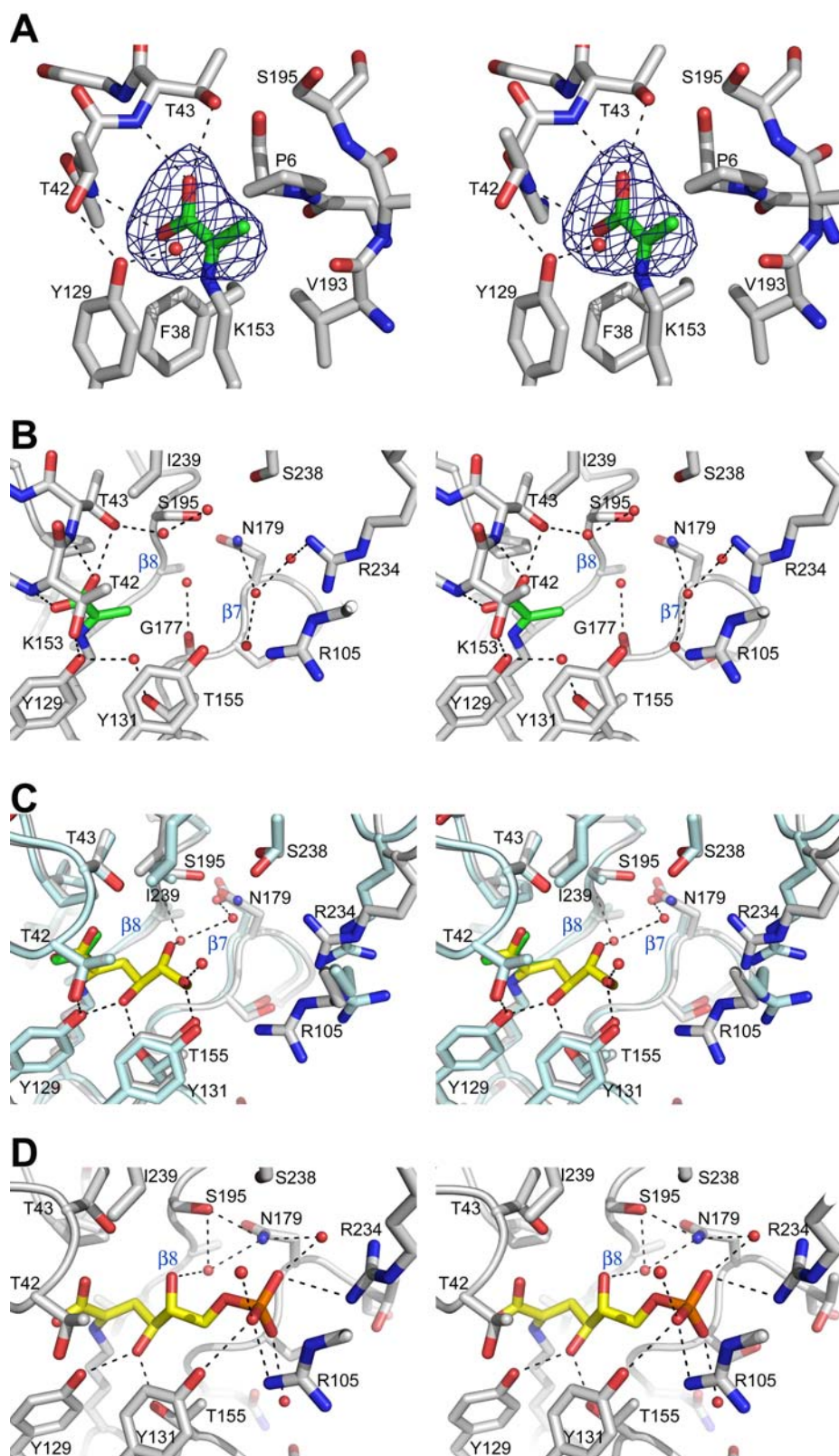


Figure 3.