Identification and characterization of *Thermoplasma acidophilum* glyceraldehyde dehydrogenase: a new class of NADP⁺-specific aldehyde dehydrogenase

Jin Hwa JUNG* and Sun Bok LEE*†¹

*Department of Chemical Engineering, Pohang University of Science and Technology, San 31, Hyoja-Dong, Pohang 790-784, Korea, and †Division of Molecular and Life Sciences, Pohang University of Science and Technology, San 31, Hyoja-Dong, Pohang 790-784, Korea

Thermoacidophilic archaea such as *Thermoplasma acidophilum* and *Sulfolobus solfataricus* are known to metabolize D-glucose via the nED (non-phosphorylated Entner–Doudoroff) pathway. In the present study, we identified and characterized a glyceraldehyde dehydrogenase involved in the downstream portion of the nED pathway. This glyceraldehyde dehydrogenase was purified from *T. acidophilum* cell extracts by sequential chromatography on DEAE-Sepharose, Q-Sepharose, Phenyl-Sepharose and Affi-Gel Blue columns. SDS/PAGE of the purified enzyme showed a molecular mass of approx. 53 kDa, whereas the molecular mass of the native protein was 215 kDa, indicating that glyceraldehyde dehydrogenase is a tetrameric protein. By MALDI–TOF–MS (matrix-assisted laser-desorption ionization–time-of-flight MS) peptide fingerprinting of the purified protein, it was found that the gene product of Ta0809 in the *T. acidophilum* genome database corresponds to the purified glyceraldehyde dehydrogenase. The

native enzyme showed the highest activity towards glyceraldehyde, but no activity towards aliphatic or aromatic aldehydes, and no activity when NAD⁺ was substituted for NADP⁺. Analysis of the amino acid sequence and enzyme inhibition studies indicated that this glyceraldehyde dehydrogenase belongs to the ALDH (aldehyde dehydrogenase) superfamily. BLAST searches showed that homologues of the Ta0809 protein are not present in the *Sulfolobus* genome. Possible differences between *T. acidophilum* (Euryarchaeota) and *S. solfataricus* (Crenarchaeota) in terms of the glycolytic pathway are thus expected.

Key words: aldehyde dehydrogenase (ALDH) superfamily, glyceraldehyde dehydrogenase, NADP⁺-specific enzyme, non-phosphorylated Entner–Doudoroff (nED) pathway, *Thermoplasma acidophilum*.

INTRODUCTION

A variety of pathways are involved in the degradation of glucose to pyruvate, and the EMP (Embden–Meyerhof–Parnas) pathway is a general route of glucose degradation in all domains of life [1]. However, in thermoacidophilic archaea, a modified version of the ED (Entner–Doudoroff) pathway appears to be more commonly used for sugar dissimilation [2–4]. It has been suggested that thermoacidophilic archaea such as *Thermoplasma acidophilum*, *Sulfolobus solfataricus* and *Thermoproteus tenax* metabolize glucose via the nED (non-phosphorylated ED) pathway (Figure 1), an ED-like pathway in which hexose intermediates are not phosphorylated. In these micro-organisms, activation via phosphorylation occurs during the latter part of the pathway, at the level of glycerate, and before merging with the EMP pathway.

The first step in the nED pathway involves the NAD(P)⁺-dependent oxidation of glucose to gluconate, catalysed by glucose dehydrogenase. Gluconate is then dehydrated by gluconate dehydratase to KDG (2-keto-3-deoxy-D-gluconate), which undergoes an aldolase cleavage to glyceraldehyde and pyruvate, catalysed by KDG aldolase. Glyceraldehyde dehydrogenase then oxidizes glyceraldehyde to glycerate, which is phosphorylated by glycerate kinase to give 2-phosphoglycerate. A second molecule of pyruvate is produced from this compound by the actions of enolase and pyruvate kinase in the EMP pathway.

All of the enzymes involved in the first part of the nED pathway, which transforms D-glucose to glyceraldehyde and pyruvate, have been identified and characterized (D-glucose dehydrogenase from *T. acidophilum* [5] and *S. solfataricus* [6], D-gluconate dehydratase from *S. solfataricus* [7,8], and KDG aldolase from *S. solfataricus* [9]). However, the genes corresponding to glyceraldehyde dehydrogenase or glycerate kinase in the downstream portion of the nED pathway have not been identified as yet.

Glyceraldehyde dehydrogenase is a key enzyme in the downstream portion of the nED pathway, and it has been confirmed experimentally that D-glyceraldehyde is specifically converted into D-glycerate by cell-free extracts of *T. acidophilum* [10]. Moreover, the entire genome of *T. acidophilum* has been sequenced [11], and four genes encoding proteins homologous to ALDH (aldehyde dehydrogenase) have been identified. However, the gene corresponding to glyceraldehyde dehydrogenase has not been identified unambiguously. In addition, the substrate specificity of glyceraldehyde dehydrogenase remains to be established experimentally.

Previously, we have described the identification and characterization of *S. solfataricus* D-gluconate dehydratase, a key enzyme of the nED pathway [7]. In the present study, we describe the identification and characterization of glyceraldehyde dehydrogenase from *T. acidophilum*, another important enzyme in the nED pathway. To our knowledge, this is the first report on

Abbreviations used: ALDH, aldehyde dehydrogenase; DTT, dithiothreitol; ED, Entner–Doudoroff; EMP, Embden–Meyerhof–Parnas; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; KDG, 2-keto-3-deoxy-D-gluconate; MALDI–TOF–MS, matrix-assisted laser-desorption ionization–time-of-flight MS; nED, non-phosphorylated ED.

¹ To whom correspondence should be addressed (email sblee@postech.ac.kr).

Figure 1 The nED pathway

the characterization of a glyceraldehyde dehydrogenase from thermoacidophilic archaea.

EXPERIMENTAL

Chemicals

D,L-Glyceraldehyde, D,L-glyceraldehyde-3-phosphate, methylglyoxal, formaldehyde, glycolaldehyde, phenylacetaldehyde, *p*-nitrobenzaldehyde, NAD⁺, NADP⁺, NADPH, glutathione, 2-mercaptoethanol and *p*-hydroxymercuribenzoic acid were purchased from Sigma-Aldrich (St. Louis, MO, U.S.A.); D-glyceraldehyde, propionaldehyde, acetaldehyde and *p*-chloromercuribenzoic acid were obtained from Fluka (St. Gallen, Switzerland), and DTT (dithiothreitol) was from USB (Cleveland, OH, U.S.A.). DEAE-Sepharose, Q-Sepharose, Phenyl-Sepharose, Sephacryl S-200 and gel filtration calibration kits were obtained from Amersham Pharmacia (Uppsala, Sweden), and Affi-Gel Blue was purchased from Bio-Rad (Hercules, CA, U.S.A.).

Strains and culture conditions

T. acidophilum (JCM 9062) was obtained from the Japan Collection of Micro-organisms. The composition of the growth medium used was (per litre): 1.0 g yeast extract, 3.0 g glucose, 1.3 g (NH₄)₂SO₄, 0.3 g KH₂PO₄, 0.5 g MgSO₄/7H₂O and 0.25 g CaCl₂/2H₂O. Final pH was adjusted to 2.0 using 10 M H₂SO₄. Cultures were grown aerobically in a 3 litre fermentor (KLF 2000, Bioengineering AG, Wald, Switzerland) at 59°C for 60 h with stirring at 50 rev./min. Growth was monitored spectrophotometrically at 600 nm.

Glyceraldehyde dehydrogenase assay

Glyceraldehyde dehydrogenase activity was determined spectrophotometrically at 340 nm. Reaction mixtures (total volume 1.5 ml), which contained 5 mM D,L-glyceraldehyde, 0.13 mM NADP⁺ as coenzyme and appropriate amounts of enzyme in 50 mM Tris/HCl (pH 8.0), were incubated for 30 min at 50°C. Mixtures were cooled on ice to stop the reaction, and any increase in absorption at 340 nm due to the reduction of NADP⁺ was monitored spectrophotometrically. Enzyme activity was calculated using the molar absorption coefficient of NADPH at 340 nm (5841 M⁻¹ · cm⁻¹). One unit of glyceraldehyde dehydrogenase activity was defined as the amount of enzyme required to reduce 1 μmol of NADP⁺ per min at 50°C and pH 8.0. All enzyme activity was determined in triplicate. Protein concentration was determined according to the method of Bradford [11a], and bovine serum albumin was used as a standard.

Purification of glyceraldehyde dehydrogenase

Cells were grown to the exponential phase and harvested by centrifugation at 4000 *g* for 20 min. Cell pellets were resuspended in Tris/HCl and crude extracts were prepared by sonication. Cell suspensions were then centrifuged at 20 000 *g* for 60 min at 4°C to remove cellular debris, and supernatants were dialysed against Tris/HCl. The dialysed supernatants were then subjected to ion-exchange chromatography using a DEAE-Sepharose column (2.5 cm × 16 cm), which had been previously equilibrated with Tris/HCl. Enzymes were eluted using a stepwise gradient of NaCl (0.1, 0.2, 0.4, 0.6 and 1.0 M) in Tris/HCl (flow rate: 1 ml/min, volume of fraction: 5 ml). Active fractions were pooled, dialysed against Tris/HCl and concentrated using a Vivaspinn concentrator membrane (Vivascience, Lincoln, U.K.).

DEAE-Sepharose active fractions were loaded onto a Q-Sepharose column (2.5 cm × 14 cm) that had been previously equilibrated with Tris/HCl. Columns were then washed with 2 volumes of buffer, and enzymes were eluted using a stepwise gradient of NaCl (0.1, 0.2, 0.4, 0.6, 0.8 and 1.0 M) in Tris/HCl (flow rate: 1 ml/min, volume of fraction: 4 ml). Active fractions were pooled, dialysed against Tris/HCl and concentrated using a concentrator membrane.

The Q-Sepharose active fraction was loaded on to a Phenyl-Sepharose column (1.0 cm × 12 cm) equilibrated with Tris/HCl containing 0.5 M ammonium sulphate. The column was then washed with Tris/HCl, and eluted with a stepwise gradient of ammonium sulphate (0.5, 0.4, 0.3, 0.2, 0.1 and 0 M) in Tris/HCl (flow rate: 1 ml/min, volume of fraction: 5 ml). Active fractions were pooled and concentrated using a concentrator membrane.

The Phenyl-Sepharose active fraction was placed on an Affi-Gel Blue column (0.5 cm × 14 cm) pre-equilibrated with Tris/HCl. The column was washed with buffer and eluted using a stepwise gradient of NaCl (0.2, 0.4, 0.4, 0.8 and 1.0 M) in Tris/HCl (flow rate: 1 ml/min, volume of fraction: 1 ml). Active fractions were pooled and concentrated using a concentrator membrane.

Molecular-mass determination

Purified D-glyceraldehyde dehydrogenase (100 μg) was applied to a Sephacryl S-200 column (2.5 cm × 100 cm) at 0.4 ml/min using a gel filtration calibration kit (Amersham Pharmacia Biotech, Sweden). The column was pre-equilibrated and eluted with Tris/HCl containing 150 mM NaCl. The molecular-mass markers used were thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), BSA (bovine serum albumin; 67 kDa), and ovalbumin (43 kDa). Eluent absorbance was monitored at 280 nm and D-glyceraldehyde dehydrogenase activity was

measured as described above. The molecular mass of native D-glyceraldehyde dehydrogenase was obtained by interpolation of a plot of log molecular-mass versus K_{av} [$K_{av} = (\text{eluate volume} - \text{void volume})/(\text{total volume} - \text{void volume})$].

Identification of purified glyceraldehyde dehydrogenase

SDS/PAGE was performed using 12 % polyacrylamide slab gels. After electrophoresis, gels were stained with Coomassie Blue. Proteins in the gels were excised, sequentially digested in-gel with trypsin, and then submitted to MALDI-TOF-MS (matrix-assisted laser-desorption ionization-time-of-flight MS). Peptide mass fingerprints were acquired using a MALDI-TOF mass spectrometer (4700 Proteomics Analyzer, Applied Biosystems). Mono-isotopic peptide mass obtained after MALDI-TOF was screened against entries in the NCBI nr protein databases using protein search programs (Mascot and ProFound). A mass accuracy of 50 p.p.m. (parts per million) for the parent ion and 0.1 Da for product ion masses were used as search parameters.

Substrate specificity and kinetic parameters

The substrate specificity of native glyceraldehyde dehydrogenase for aldehyde groups was determined using NAD^+ or NADP^+ as a cofactor. The aldehyde substrates used in this study were as follows: D-glyceraldehyde, D,L-glyceraldehyde, D,L-glyceraldehyde-3-phosphate, glycolaldehyde, methylglyoxal, formaldehyde, acetaldehyde, propionaldehyde, phenylacetaldehyde and *p*-nitrobenzaldehyde. The kinetic parameters for glyceraldehyde dehydrogenase were determined by measuring its enzyme activity at various D,L-glyceraldehyde and NADP^+ concentrations. Apparent $V_{\max}$ and K_m values were calculated by fitting initial rate data to the Michaelis–Menten equation using a non-linear regression analysis program (Sigma Plot 7.0).

Optimum temperature, thermostability, optimum pH, and pH stability

The effect of temperature and pH on the enzymatic activity of glyceraldehyde dehydrogenase was determined using the assay method described above. Temperature profiles were determined between 30 and 90 °C, and thermostability was determined at 50, 60, 70 and 80 °C by incubating the native enzyme at 100 $\mu\text{g/ml}$ in 50 mM Tris/HCl (pH 8.0). At appropriate times, incubated enzyme solutions were immediately cooled on ice, and residual enzyme activity was determined. The effect of pH on native glyceraldehyde dehydrogenase activity was determined at 50 °C in 50 mM potassium phosphate buffer (pH 6–8), 50 mM Tris/HCl (pH 7–9) and 50 mM glycine/NaOH buffer (pH 9–11). The pH stability was examined by pre-incubating enzyme solutions at various pHs for 24 h and assaying residual enzyme activity at 50 °C and pH 8.0.

Effect of divalent ions and thiol reagents

The effect of divalent ions (Ba^{2+} , Ca^{2+} , Cd^{2+} , Cu^{2+} , Hg^{2+} , Mg^{2+} , Mn^{2+} , Ni^{2+} and Zn^{2+}) and EDTA on glyceraldehyde dehydrogenase was determined by measuring its activity after incubating it with 50 mM Tris/HCl (pH 8.0) at 4 °C for 12 h in the presence of 1 mM of each of the divalent ions or EDTA. The effect of thiol compounds (glutathione, DTT and 2-mercaptoethanol) and thiol-modifying reagents (*p*-chloromercuribenzoic acid and *p*-hydroxymecuribenzoic acid) on enzymatic activity was determined using the same procedure, but using these thiol compounds at 1 mM instead of divalent ions.

Table 1 Purification of glyceraldehyde dehydrogenase

Enzyme activity was assayed by measuring the amount of NADPH produced from NADP^+ at 50 °C using a standard assay method. One unit of glyceraldehyde dehydrogenase activity was defined as the amount of enzyme required to reduce 1 μmol of NADP^+ per min.

Purification step	Total protein (mg)	Total activity (unit)	Specific activity (units/mg)	Purification (fold change)
Crude extract	1460.8	131.5	0.09	1
DEAE Sepharose	81.0	19.4	0.24	2.7
Q Sepharose	36.8	9.4	0.26	2.9
Phenyl Sepharose	9.7	2.9	0.30	3.3
Affi-Gel Blue	0.77	2.2	2.85	31.7

Table 2 Substrate specificity of glyceraldehyde dehydrogenase

Enzyme activity was assayed by determining the amount of NADPH or NADH produced by 5 mM of each aldehyde substrate in 30 min at 50 °C.

Substrate	Relative activity (%)	
	NADP^+	NAD^+
D,L-glyceraldehyde	100	0
D-glyceraldehyde	96	0
D,L-glyceraldehyde-3-phosphate	33	0
Glycolaldehyde	14	0
Methylglyoxal	0	0
Formaldehyde	0	0
Acetaldehyde	0	0
Propionaldehyde	0	0
Phenylacetaldehyde	0	0
<i>p</i> -nitrobenzaldehyde	0	0

RESULTS

Purification of glyceraldehyde dehydrogenase

The purification procedure is summarized in Table 1. After sequential chromatography on DEAE-Sepharose, Q-Sepharose, Phenyl-Sepharose and Affi-Gel Blue columns, glyceraldehyde dehydrogenase was purified 32-fold. The most effective step in this enrichment procedure was Affi-Gel Blue chromatography, which increased the specific activity by 9.6-fold. The glyceraldehyde dehydrogenase activity was not detected in other fractions under our assay conditions, indicating that purified glyceraldehyde dehydrogenase is the sole enzyme responsible for the dehydrogenation of glyceraldehyde *in vivo*.

SDS/PAGE of samples obtained from different steps in the purification process was performed in 12 % polyacrylamide gels. The molecular mass of the denatured enzyme, determined by SDS/PAGE, was approx. 53 kDa, whereas the molecular mass of the native enzyme was estimated to be 215 kDa using a calibrated Sephacryl S-200 column (results not shown). These results indicate that the *T. acidophilum* glyceraldehyde dehydrogenase in its native conformation is a homotetramer.

Catalytic properties of glyceraldehyde dehydrogenase

The enzymatic activity and substrate specificity of glyceraldehyde dehydrogenase were investigated using various aldehyde substrates (Table 2). Of the substrates tested, the purified enzyme showed the highest activity towards glyceraldehyde (D-glyceraldehyde and D,L-glyceraldehyde). It also catalysed the oxidation of D-glyceraldehyde-3-phosphate and glycolaldehyde. However, its activity towards D-glyceraldehyde-3-phosphate or glycolaldehyde was only 33 % and 14 % respectively of that

Figure 2 Effect of temperature on the activity (A) and stability (B) of glyceraldehyde dehydrogenase from *T. acidophilum*

Assays were performed as described in the Experimental section at 50°C (■), 60°C (▲), 70°C (▼) and 80°C (●).

towards D,L-glyceraldehyde. On the other hand, the enzyme displayed no activity towards ketoaldehyde (methylglyoxal), alkyl aldehydes (formaldehyde, acetaldehyde and propionaldehyde) or aromatic aldehydes (phenylacetaldehyde and *p*-nitrobenzaldehyde). In addition, it only showed activity when NADP⁺ was used as a cofactor. Regardless of the substrates used, the enzyme was completely inactive when NAD⁺ was substituted for NADP⁺.

The effect of substrate concentration on the activity of purified glyceraldehyde dehydrogenase activity was examined in the range 0–3 mM for glyceraldehyde and 0–0.15 mM for NADP⁺. Its K_m and V_{max} values for D,L-glyceraldehyde were 0.33 mM and 0.30 unit/mg, and for NADP⁺ were 0.017 mM and 0.30 unit/mg respectively.

The effect of temperature and pH on glyceraldehyde dehydrogenase activity are shown in Figures 2 and 3 respectively. The purified enzyme showed optimal activity between 50 and 55°C (Figure 2A), which is slightly lower than the optimal temperature required for the growth of *T. acidophilum* (55–60°C). Its thermostability was also measured at 50, 60, 70 and 80°C. At 60°C, glyceraldehyde dehydrogenase was stable for 2 h, but at 80°C its activity was decreased to less than 50% within 20 min (Figure 2B). Its optimal pH was approx. 8.0 (Figure 3), which is somewhat higher than the intracellular pH of *T. acidophilum* [12]. Its pH stability was also determined by pre-incubating the enzyme at various pHs for 24 h. It was found to be stable between pH 6.0 and 8.0.

The effect of metal ions on enzyme activity was also examined. Of the divalent metal ions tested, Ca²⁺, Cd²⁺, Cu²⁺ and Hg²⁺

Figure 3 Effect of pH on the activity of glyceraldehyde dehydrogenase from *T. acidophilum*

Assays were performed as described in the Experimental section in 50 mM citrate-NaOH buffer (▲), 50 mM potassium phosphate buffer (■), 50 mM Tris/HCl (▼) and 50 mM glycine/NaOH (●).

Table 3 Effect of divalent ions on glyceraldehyde dehydrogenase

Purified enzyme was pre-incubated with each metal solution (1 mM) for 12 h at 4°C. Enzyme activity was assayed as described in the Experimental section.

Divalent ion	Relative activity (%)
None	100
Ba ²⁺	102
Mg ²⁺	76
Mn ²⁺	56
Zn ²⁺	54
Ni ²⁺	42
Cd ²⁺	6.7
Hg ²⁺	0.7
Cu ²⁺	0.4
Ca ²⁺	< 0.1
EDTA	93

Table 4 Effect of thiol compounds and thiol-modifying reagents on the activity of glyceraldehyde dehydrogenase

Purified enzyme was pre-incubated with each compound (at 1 mM) for 12 h at 4°C. Enzyme activity was assayed as described in the Experimental section.

Compound	Relative activity (%)
None	100
Glutathione	218
DTT	216
2-mercaptoethanol	211
<i>p</i> -chloromercuribenzoic acid	0
<i>p</i> -hydroxymercuribenzoic acid	0

were found to significantly inhibit enzyme activity (Table 3), and Ca²⁺ was the most potent inhibitor, with a 50% inhibitory concentration of approx. 20 μM.

Table 4 shows the effect of various thiols on the activity of glyceraldehyde dehydrogenase. In the presence of 1 mM glutathione, DTT, or 2-mercaptoethanol, enzyme activity was enhanced more than 2-fold. On the other hand, thiol-modifying reagents such as *p*-chloromercuribenzoic acid or *p*-hydroxymercuribenzoic acid completely inhibited its activity. The results shown in Table 4 indicate that a free SH group is required for the catalytic activity of glyceraldehyde dehydrogenase.

Figure 4 Amino acid sequence alignment of *T. acidophilum* glyceraldehyde dehydrogenase orthologues

The sequences aligned are: Tac, *T. acidophilum* (GenBank® accession number CAC11938); Tvo, *T. volcanium* (GenBank® accession number BAB60192); Pto, *P. torridus* (GenBank® accession number YP_023110) and Fac, *F. acidarmanus* (GenBank® accession number ZP_00307139). Invariant residues are indicated by dark grey. Putative active sites common to the ALDH superfamily are indicated by arrows. The 16 residues conserved in the ALDH superfamily are highlighted in black with white letters. The sequences of the 10 most conserved motifs in the ALDH superfamily are shown at the bottom. The glycine motif involved in NAD(P)⁺ binding, G-X-G-X-X-G, is also shown (between residues 203 and 209 of Ta0809). X denotes positions without a clear residue consensus.

Identification of the gene encoding glyceraldehyde dehydrogenase

The protein spot corresponding to glyceraldehyde dehydrogenase was excised from an SDS/PAGE gel, digested in-gel with trypsin, and then analysed by MALDI-TOF-MS. When the mono-isotopic peptide mass obtained via MALDI-TOF was screened against the NCBI protein database, its mass spectrum matched the sequence of Ta0809 in the *T. acidophilum* genome database (Mascot score = 300, sequence coverage = 63 %). Moreover, the molecular-mass deduced from the amino acid sequence (54 782 Da) was in agreement with the molecular mass of purified glyceraldehyde dehydrogenase in a denaturing gel (53 kDa). Consequently, we confirmed that the protein purified from *T. acidophilum* was glyceraldehyde dehydrogenase and that its open reading frame corresponds to that of Ta0809. We named the corresponding gene *gld*, as it encodes glyceraldehyde dehydrogenase.

Sequence comparison with ALDHs

A BLAST search of the available sequence databases suggested that the deduced amino acid sequence of *T. acidophilum* glyceraldehyde dehydrogenase (Ta0809) is homologous to that of several NAD(P)⁺-dependent ALDHs. This search showed that the amino acid sequence of Ta0809 is highly homologous with that of proteins in other thermoacidophilic archaea such as *Thermoplasma volcanium*, *Picrophilus torridus* and *Ferroplasma acidarmanus* (Figure 4). The calculated sequence identity was 85 % for the putative ALDH of *T. volcanium* (GenBank® accession number BAB60192), 71 % for the putative succinate-semialdehyde dehydrogenase of *P. torridus* (YP_023110), and 71 % for the NAD-dependent ALDH of *F. acidarmanus* (ZP_00307139). All of these micro-organisms belong to the Euryarchaeota. However, no homologues of the Ta0809 protein could be found in the *S. solfataricus* [4], *Sulfolobus tokodaii*

[13] and *Sulfolobus acidocaldarius* [14] genome databases, which belong to the Crenarchaeota. This result was unexpected, because it has been postulated that *T. acidophilum* and *S. solfataricus* share the same nED pathway [1–4].

Further analysis of the protein sequence of Ta0809 indicated that it belongs to the ALDH superfamily. The 16 amino acid residues that are found in more than 95 % of ALDHs [15,16] were all conserved (highlighted in black with white letters in Figure 4), and included Glu²⁴⁷ and Cys²⁸¹, which have been previously implicated in the catalytic activity of ALDHs [15,16]. A glycine motif involved in NAD(P)⁺ binding, G-X-G-X-X-G (GRGSEIG), was also found between residues 203 and 209. This motif has been found in several ALDHs capable of catalysing irreversible dehydrogenation reactions. Of the 10 motifs most conserved in ALDHs [16], motif 10 is the least conserved in glyceraldehyde dehydrogenase (see Figure 4). It appears that this motif differentiates glyceraldehyde dehydrogenase from ALDHs.

DISCUSSION

In the present study, we purified glyceraldehyde dehydrogenase from cell extracts of the thermoacidophilic archaeon *T. acidophilum* and identified the gene that encodes this enzyme in the *T. acidophilum* genomic database. Although its activity in *T. acidophilum* has been known for some time [10], the nucleotide and amino acid sequences encoding the glyceraldehyde dehydrogenase enzyme have not been reported, which is why the corresponding glyceraldehyde dehydrogenase gene could not be discovered in genomic data published to date. By MALDI-TOF-MS peptide mass fingerprinting of the purified glyceraldehyde dehydrogenase, we were able to identify the gene (Ta0809; *gld*) in the *T. acidophilum* genome database. Previously, Verhees et al. [2] predicted Ta0809 to be a NAD⁺-specific

Table 5 Classification of aldehyde dehydrogenase superfamily

Adapted from Perozich et al. [16], except glyceraldehyde dehydrogenase. DH, dehydrogenase.

Family	Abbreviation	Specificity	Quaternary structure	Coenzyme(s)
'Class 3' trunk				
Betaine ALDH	BALDH	Specific	Tetramer	NAD
'Amino acid intermediate' sub-branch				
Methyl malonyl semialdehyde DH	MMSALDH	Specific	Tetramer	NAD and CoA
γ -Glutamyl semialdehyde DH	GGSALDH	Specific	Dimer	NAD
Turgor-responsive ALDH	Turgor ALDH	—	—	—
'Class 3' sub-branch				
Class 3 ALDH		Variable	Dimer	NAD or NADP
Aromatic-metabolizing ALDH	Aromatic ALDH	Specific	Tetramer	NAD
Succinic semialdehyde DH	SSALDH	Specific	Tetramer	NAD or NADP
Glyceraldehyde-3-phosphate DH (non-phosphorylating)	GAPDH	Specific	Tetramer	NADP
Glyceraldehyde DH	GADH	Specific	Tetramer	NADP
'Class 1/2' trunk				
Class 1 ALDH		Variable	Tetramer	NAD
Class 2 ALDH		Variable	Tetramer	NAD
Fungal ALDH		Variable	Tetramer	NAD
10-Formyltetrahydrofolate DH	FTDH	Specific	Tetramer	NADP
2-Hydroxymuconic semialdehyde DH	HMSALDH	Specific	Dimer	NAD

GAPDH (glyceraldehyde-3-phosphate dehydrogenase), and with less confidence, an NAD(P)⁺-specific ALDH. BLAST searches of the NCBI database showed that other thermoacidophilic archaea such as *T. volcanium*, *P. torridus* and *F. acidarmanus* also possess this protein and that their amino acid sequences are highly homologous (71–85 % identity) to Ta0809.

An analysis of the amino acid sequence of Ta0809 and inhibition studies indicated that the glyceraldehyde dehydrogenase identified in this study belongs to the ALDH superfamily. Sequence alignment enabled us to identify the conserved residues commonly found in the ALDH superfamily enzymes. It is known that ALDH superfamily proteins have typical conserved sites e.g. 4 amino acid residues are invariant and 12 residues are found in more than 95 % of the sequences [16]. All of these 16 residues are conserved in Ta0809 (see Figure 4) and include Glu²⁴⁷ and Cys²⁸¹, which have been implicated in the catalytic activity of ALDHs [15,16]. These results suggest that the catalytic mechanism of glyceraldehyde dehydrogenase resembles that of ALDHs.

The results of inhibition studies also support this postulation. The effect of thiol reagents on the enzymatic activity of Ta0809 appears to be similar to their effects on other ALDHs. Moreover, Ta0809 is activated by the presence of thiol compounds such as glutathione, DTT and 2-mercaptoethanol (see Table 4). It has been reported that thiol compounds cause either an increase or no appreciable change in the activity of mammalian ALDHs [17,18]. On the other hand, thiol-modifying reagents such as *p*-chloromercuribenzoic acid or *p*-hydroxymercuribenzoic acid completely inhibited the enzymatic activity of Ta0809, and a similar inhibitory effect of thiol-modifying reagents was also reported for ALDH purified from bovine lens [17]. Since *p*-chloromercuribenzoic acid and *p*-hydroxymercuribenzoic acid specifically modify the cysteine residues of proteins, it is probable that Cys²⁸¹ is involved in the catalytic activity of glyceraldehyde dehydrogenase.

It appears that the glyceraldehyde dehydrogenases of thermoacidophilic archaea form a distinct family within the superfamily of ALDHs. Glyceraldehyde dehydrogenase only uses NADP⁺ as a cofactor. Of the 13 ALDH families, most ALDHs are NAD⁺-specific, whereas two families, 10-formyltetrahydrofolate dehydrogenases and non-phosphorylating GAPDHs, are NADP⁺-specific (Table 5). Interestingly, 15 of the Ta0809 amino acid

sequences in the cofactor-binding GXGXXXG region (ITGRGS-EIGDYIVEH) are identical to those of NADP⁺-specific GAPDH from *Streptomyces mutans* (JMB1999) [19]. The majority of ALDHs act on a wide variety of aliphatic or aromatic aldehydes; however, glyceraldehyde dehydrogenase preferentially oxidizes glyceraldehyde. This enzyme also shows some activity towards other hydroxyaldehydes such as glyceraldehyde-3-phosphate and glycolaldehyde, but no activity towards alkyl or aromatic aldehydes.

The structure of several members of the ALDH superfamily is known. These include human mitochondrial aldehyde dehydrogenase [20], betaine aldehyde dehydrogenase from cod liver [21], rat retinal dehydrogenase [22], fatty aldehyde dehydrogenase from the luminescent bacterium *Vibrio harveyi* [23], GAPDH from *Streptococcus mutans* [19], and medium-chain aldehyde dehydrogenase from *Escherichia coli* [24]. Alignment of the sequences of the ALDH superfamily proteins indicated that the motif-10 region is probably responsible for co-ordinating with the substrate molecules required for glyceraldehyde dehydrogenase activity. In this regard, it is noteworthy that motif-10 in ALDHs is least conserved in glyceraldehyde dehydrogenase (see Figure 4). However, the amino acid sequence in the region corresponding to motif-10 is well conserved within thermoacidophilic archaea. This would account for the observed differences in substrate specificity between glyceraldehyde dehydrogenase and other ALDHs.

As described above, BLAST searches showed that the amino acid sequence of *gld* is highly homologous to proteins in the Euryarchaeota, such as *T. volcanium*, *P. torridus* and *F. acidarmanus*. However, no homologues were found in the Crenarchaeota, such as *S. solfataricus*, *S. tokodaii* and *S. acidocaldarius*. This result was unexpected because it has been postulated that *T. acidophilum* and *S. solfataricus* share the same nED pathway [1–4]. Recently, we compared the activity of the glyceraldehyde dehydrogenase of *T. acidophilum* with that of *S. solfataricus* using crude cell extracts [25]. *S. solfataricus* showed no appreciable glyceraldehyde dehydrogenase activity whereas *T. acidophilum* did. We confirmed the absence of NADP⁺-dependent glyceraldehyde dehydrogenase activity in *S. solfataricus* by repeating the same experiment three times. When NAD⁺ was used instead of NADP⁺ as a cofactor, no glyceraldehyde

dehydrogenase activity was observed in either *T. acidophilum* or *S. solfataricus*. On the other hand, their GAPDH activity level was similar.

The absence of NAD(P)⁺-dependent glyceraldehyde dehydrogenase activity in *S. solfataricus* cell-free extracts and the absence of homologues of the glyceraldehyde dehydrogenase gene in the *Sulfolobus* genome database suggest that the conversion of glyceraldehyde to glycerate in the Crenarchaeota is catalysed by an enzyme other than glyceraldehyde dehydrogenase. In this regard, it is noteworthy that the molybdenum-containing aldehyde oxidoreductase of *S. acidocaldarius* exhibits oxidizing activity towards glyceraldehyde, glyceraldehyde-3-phosphate, isobutyraldehyde, formaldehyde, acetaldehyde and propionaldehyde [26]. At the optimum pH of this enzyme (6.7), which is close to the intracellular pH of *Sulfolobus*, glyceraldehyde-oxidizing activity is predominant. This suggests that molybdenum-containing aldehyde oxidoreductase may function as a catalyst for glyceraldehyde oxidation in *S. acidocaldarius*. This protein has an apparent molecular mass of 177 kDa and consists of three subunits of 80.5 kDa (α), 32 kDa (β) and 19.5 kDa (γ). The N-terminal amino acid sequence of these proteins corresponds to the Saci.2271, Saci.2269 and Saci.2270 genes in the *S. acidocaldarius* genome [14]. Using BLAST searches, we were also able to identify proteins homologous to *S. acidocaldarius* aldehyde oxidoreductase in the genome of other thermoacidophilic archaea such as *S. solfataricus*, *S. tokodaii* and *T. tenax*. Interestingly, however, no homologues could be found in the genome databases of *T. acidophilum* and *T. volcanium*, which may reflect differences in the glyceraldehyde oxidation mechanism in thermoacidophilic archaea.

In summary, the glyceraldehyde dehydrogenase of *T. acidophilum* was found to have strict specificity towards NADP⁺ and glyceraldehyde. Biochemical data and amino acid sequence comparisons indicate that this glyceraldehyde dehydrogenase belongs to a new class of the ALDH superfamily. It was also found that the gene corresponding to glyceraldehyde dehydrogenase (*gld*) is Ta0809 in the *T. acidophilum* genome database. Importantly, the results of the present study indicate that enzyme(s) involved in glyceraldehyde oxidation are probably different between Euryarchaeota (e.g. *T. acidophilum*) and Crenarchaeota (e.g. *S. solfataricus*). This study represents a first step towards understanding the glyceraldehyde dehydrogenase involved in the glycolysis of extreme thermoacidophilic archaea via nonphosphorylated intermediates. Further studies on the structure and the catalytic mechanism of glyceraldehyde dehydrogenase are needed for a complete understanding of the mechanistics and the *in vivo* role of this enzyme in thermoacidophilic archaea. Recently, Ahmed et al. [27] provided evidence for a semi-phosphorylated variant of the ED pathway in Crenarchaeota (*S. solfataricus* and *T. tenax*). Interestingly, the key enzyme involved in phosphorylation of the KDG intermediate, KDG kinase, has also been found in *T. acidophilum* [28]. This indicates that a semi-phosphorylated ED bypass is also present in Euryarchaeota. In view of this, it is possible that the enzymatic activity of glyceraldehyde dehydrogenase is different depending on the demand for phosphorylated and non-phosphorylated intermediates in the cell. We are currently exploring this possibility.

This work was supported by a grant from the 21C Frontier Microbial Genomics and Applications Center Program.

REFERENCES

- Romano, A. H. and Conway, T. (1996) Evolution of carbohydrate metabolic pathways. *Res. Microbiol.* **147**, 448–455
- Verhees, C. H., Kengen, S. W. M., Tuininga, J. E., Schut, G. J., Adams, M. W. W., De Vos, W. M. and Van Der Oost, J. (2003) The unique features of glycolytic pathways in archaea. *Biochem. J.* **377**, 819–822
- Lamble, 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
- She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C. Y., Clausen, I. G., Curtis, B. A., DeMoors, A. et al. (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7835–7840
- Smith, L. D., Budgen, N., Bungard, S. J., Danson, M. J. and Hough, D. W. (1989) Purification and characterization of glucose dehydrogenase from the thermoacidophilic archaeobacterium *Thermoplasma acidophilum*. *Biochem. J.* **261**, 973–977
- Giardina, P., De Biasi, M. G., De Rosa, M., Gambacorta, A. and Buonocore, V. (1986) Glucose dehydrogenase from the thermoacidophilic archaeobacterium *Sulfolobus solfataricus*. *Biochem. J.* **239**, 517–522
- Kim, S. and Lee, S. B. (2005) Identification and characterization of *Sulfolobus solfataricus* D-gluconate dehydratase: a key enzyme in the nonphosphorylated Entner–Doudoroff pathway. *Biochem. J.* **387**, 271–280
- Lamble, H. J., Milburn, C. C., Taylor, G. L., Hough, D. W. and Danson, M. J. (2004) Gluconate dehydratase from the promiscuous Entner–Doudoroff pathway in *Sulfolobus solfataricus*. *FEBS Lett.* **576**, 133–136
- 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
- Budgen, N. and Danson, M. J. (1986) Metabolism of glucose via a modified Entner–Doudoroff pathway in the thermoacidophilic archaeobacterium *Thermoplasma acidophilum*. *FEBS Lett.* **196**, 207–210
- Ruepp, A., Graml, W., Santos-Martinez, M. L., Koretke, K. K., Volker, C., Mewes, H. W., Frishman, D., Stocker, S., Lupas, A. N. and Baumeister, W. (2000) The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. *Nature (London)* **407**, 508–513
- Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254
- Searcy, D. G. (1976) *Thermoplasma acidophilum*: Intracellular pH and potassium concentration. *Biochim. Biophys. Acta* **451**, 278–286
- Kawarabayashi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankaï, A., Kosugi, H., Hosoyama, A. et al. (2005) Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain 7. *DNA Res.* **8**, 123–140
- Chen, L., Brugger, K., Skovgaard, M., Redder, P., She, Q., Torarinnsson, E., Greve, B., Awayez, M., Zibat, A., Klenk, H. P. and Garrett, R. A. (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. *J. Bacteriol.* **187**, 4992–4999
- Hempel, J., Nicholas, H. and Lindahl, R. (1993) Aldehyde dehydrogenases: widespread structural and functional diversity within a shared framework. *Protein Sci.* **2**, 1890–1900
- Perozich, J., Nicholas, H., Wang, B. C., Lindahl, R. and Hempel, J. (1999) Relationships within the aldehyde dehydrogenase extended family. *Protein Sci.* **8**, 137–146
- Ting, H. H. and Crabbe, M. J. (1983) Bovine lens aldehyde dehydrogenase: kinetics and mechanism. *Biochem. J.* **215**, 361–368
- Bedino, S., Testore, G. and Obert, F. (1990) Initial characterization of aldehyde dehydrogenase from rat testis cytosol. *Biol. Chem. Hoppe Seyler* **371**, 95–101
- Cobessi, D., Tete-Favier, F., Marchal, S., Azza, S., Branlant, G. and Aubry, A. (1999) Apo and holo crystal structures of an NADP-dependent aldehyde dehydrogenase from *Streptococcus mutans*. *J. Mol. Biol.* **290**, 161–173
- Steinmetz, C. G., Xie, P., Weiner, H. and Hurley, T. D. (1997) Structure of mitochondrial aldehyde dehydrogenase: the genetic component of ethanol aversion. *Structure* **5**, 701–711
- Johansson, K., El-Ahmad, M., Ramaswamy, S., Hjelmqvist, L., Jorvall, H. and Eklund, H. (1998) Structure of betaine aldehyde dehydrogenase at 2.1 Å resolution. *Protein Sci.* **7**, 2106–2117
- Lamb, A. L. and Newcomer, M. E. (1999) The structure of retinal dehydrogenase type II at 2.7 Å resolution: implications for retinal specificity. *Biochemistry* **38**, 6003–6011
- Ahvazi, B., Coulombe, R., Delarge, M., Vedadi, M., Zhang, L., Meighen, E. and Vrielink, A. (2000) Crystal structure of the NADP⁺-dependent aldehyde dehydrogenase from *Vibrio Harveyi*: structural implications for cofactor specificity and affinity. *Biochem. J.* **349**, 853–861
- Gruez, A., Roig-Zamboni, V., Grisel, S., Salomoni, A., Valencia, C., Campanacci, V., Tegoni, M. and Cambillau, C. (2004) Crystal structure and kinetics identify *Escherichia coli* YdcW gene product as a medium-chain aldehyde dehydrogenase. *J. Mol. Biol.* **343**, 29–41

-
- 25 Noh, M. (2005) Identification and characterization of glycerate kinase from the thermoacidophilic archaeon *Thermoplasma acidophilum*. Master Thesis, Pohang University of Science and Technology
- 26 Kardinahl, S., Schmidt, C. L., Hansen, T., Anemuller, S., Petersen, A. and Schafer, G. (1999) The strict molybdate-dependence of glucose-degradation by the thermoacidophilic *Sulfolobus acidocaldarius* reveals the first crenarchaeotic molybdenum containing enzyme – an aldehyde oxidoreductase. *Eur. J. Biochem.* **260**, 540–548
- 27 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
- 28 Jung, J. H. and Lee, S. B. (2005) Identification and characterization of *Thermoplasma acidophilum* 2-keto-3-deoxy-D-gluconate kinase: a novel enzyme essential for the partially non-phosphorylated Entner–Doudoroff pathway. *Biotechnol. Bioprocess Eng.* **10**, 535–539
-

Received 1 November 2005/13 March 2006; accepted 27 March 2006

Published as BJ Immediate Publication 27 March 2006, doi:10.1042/BJ20051763