

# The glyceraldehyde-3-phosphate dehydrogenase of *Clostridium acetobutylicum*: isolation and purification of the enzyme, and sequencing and localization of the *gap* gene within a cluster of other glycolytic genes

Wiebke Schreiber and Peter Dürre

Author for correspondence: Peter Dürre. Tel: +49 731 50 22710. Fax: +49 731 50 22719.  
e-mail: peter.duerre@biologie.uni-ulm.de

Mikrobiologie und  
Biotechnologie, Universität  
Ulm, D-89069 Ulm,  
Germany

**Glyceraldehyde-3-phosphate dehydrogenase was purified from *Clostridium acetobutylicum* by sequential ammonium sulfate precipitation, gel filtration and anion-exchange chromatography (to a specific activity of 27 U mg<sup>-1</sup>). The enzyme had a molecular mass of 40 kDa as determined by SDS-PAGE and a native molecular mass of 160 kDa as determined by nondenaturing PAGE, indicating that it has a homotetrameric composition. Its pH optimum was between 8.5 and 9.3. The corresponding gene (*gap*) was cloned and sequenced from *C. acetobutylicum* DSM 792 and found to cluster with other genes of enzymes from the glycolytic pathway (*pgk*, phosphoglycerate kinase; *tpi*, triosephosphate isomerase; *pgm(i)*, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase). No sequences resembling rho-independent transcription terminators were found in the intergenic regions. A plasmid carrying the clostridial *gap* gene complemented an *Escherichia coli gap* mutant.**

Keywords: GAPDH, glycolytic gene cluster, *Clostridium acetobutylicum*

## INTRODUCTION

*Clostridium acetobutylicum* is an obligately anaerobic, Gram-positive, spore-forming bacterium which is characterized by a biphasic fermentation pattern. A typical butyric acid fermentation is carried out during the exponential-growth phase with acetate and butyrate as main fermentation products from starch or glucose. At the end of exponential growth, a metabolic transition to the production of organic solvents (acetone, butanol) takes place (Jones & Woods, 1986; Dürre, 1998). During the solventogenic growth phase, glycolytic reactions are most important for energy conservation in this organism, leading to a net production of two moles of ATP per mole of hexose consumed. In contrast to genes of

enzymes involved in solvento- and acidogenesis, only limited information is available about the structure and organization of genes encoding glycolytic enzymes in *C. acetobutylicum*. So far, only the *pfk* (phosphofructokinase) gene has been cloned and sequenced. Adjacent to *pfk*, the pyruvate kinase gene (*pyk*) has been identified and it was shown that the two genes are organized within a single operon (Belouski *et al.*, 1998). One key enzyme controlling glycolytic activity is the NAD<sup>+</sup>-dependent glyceraldehyde-3-phosphate dehydrogenase (GAPDH, EC 1.2.1.12) catalysing the reversible phosphorylation of glyceraldehyde 3-phosphate to 1,3-bisphosphoglycerate. GAPDH is inhibited by a high intracellular NADH/NAD<sup>+</sup> ratio (Girbal & Soucaille, 1994), which would otherwise favour the biotechnologically desired butanol production.

In this study, the purification of the clostridial NAD<sup>+</sup>-dependent GAPDH is presented. Cloning of the respective gene revealed a clustering with further genes (*tpi*, *pgk*, *pgm(i)*) encoding enzymes catalysing sequential reactions in the Embden–Meyerhof–Parnas pathway of glycolysis. Detailed knowledge of GAPDH and its

**Abbreviations:** GAPDH, glyceraldehyde-3-phosphate dehydrogenase; PGK, phosphoglycerate kinase; PGM, phosphoglycerate mutase; PGM(i), 2,3-bisphosphoglycerate-independent phosphoglycerate mutase; TPI, triosephosphate isomerase.

The GenBank accession number for the sequence reported in this paper is AF043386.

corresponding gene will offer the opportunity of constructing tailor-made strains that allow high glycolytic flux at high NADH concentrations and thus enhanced butanol production.

## METHODS

**Bacterial strains, plasmids and growth conditions.** *C. acetobutylicum* DSM 792 (Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany) was grown under strictly anaerobic conditions at 37 °C in clostridial basal medium (O'Brien & Morris, 1971). Continuous culture experiments were performed as described by Bahl *et al.* (1982) and were carried out in a Biostat B fermenter (Braun) with a working volume of 8 l. *Clostridium pasteurianum* DSM 525 was grown under anaerobic conditions in yeast/glucose medium (101 mM glucose, 1% yeast extract, 200 mM CaCO<sub>3</sub>). *Escherichia coli* was cultivated aerobically at 37 °C in LB medium (Sambrook *et al.*, 1989) supplemented with ampicillin (100 µg ml<sup>-1</sup>) when required.

Functional complementation experiments were performed with the *E. coli gap* mutant W3CG (Ganter & Plückthun, 1990), in which this gene was inactivated by transposon Tn10 insertion. The mutant was cultured at 37 °C in M63 medium (Silhavy *et al.*, 1984) supplemented with malate or succinate (0.4%, w/v), glycerol (0.1%, v/v) and tetracycline (20 µg ml<sup>-1</sup>). In screening experiments for functional complementation, bacteria were grown in M63 medium plus glucose or LB medium supplemented with tetracycline and ampicillin.

**Preparation of cell extracts.** Anaerobic conditions were maintained throughout the whole procedure. Culture aliquots (6 l) were removed from the continuous culture of *C. acetobutylicum* under sterile and strictly anaerobic conditions, employing tightly sealing centrifuge bottles and an anaerobic cabinet (Mecaplex). After harvesting, the cells were washed once in 1/10 0.1 M potassium phosphate buffer with 5 mM 2-mercaptoethanol (pH 7.4). The cell pellet was suspended in 1/100 buffer, supplemented with DNase I (0.1 mg ml<sup>-1</sup>) and incubated for 30–60 min at 37 °C. Lysis was achieved by 3–5 passages through a French pressure cell at 170 MPa. Cell debris was removed by centrifugation at 40000 g for 1 h at 4 °C. The clear supernatant (crude extract) was used immediately or stored at –70 °C.

**Enzyme assays.** The activity of GAPDH was determined spectrophotometrically at 30 °C by monitoring the generation of NADH (physiological direction, oxidation of glyceraldehyde 3-phosphate in the presence of sodium arsenate according to Ganter & Plückthun (1990) at 365 nm. A coupled assay in which aldolase cleaves fructose 1,6-bisphosphate to dihydroxyacetone phosphate and glyceraldehyde 3-phosphate and the latter product serves as actual substrate, was usually used during enzyme purification (Lovitt *et al.*, 1988). One unit of enzyme activity was defined as the amount of enzyme required to form 1 µmol NADH min<sup>-1</sup>.

**Purification of GAPDH.** If not otherwise noted, all steps were performed at 4–8 °C. As a first purification step, a 50–90% (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> fractionation was carried out. The precipitated material was dissolved in 20 mM BTP (Bistrispropane) buffer (pH 6.5) containing 1 mM EDTA, 2 mM DTT and 5% glycerol and applied onto a gel filtration column (Superdex 200 preparatory grade; Amersham Pharmacia Biotech) equilibrated with the same buffer. The active fractions were pooled and concentrated if necessary using Diaflo chambers and YM or PM membranes (Amicon). The protein solution was then

applied to a Q-Sepharose HP column (C26/60; Amersham Pharmacia Biotech) equilibrated with BTP buffer. After the column had been washed with the same buffer, GAPDH was eluted with 5 column volumes of a linearly increasing NaCl gradient from 0 to 0.25 M NaCl in BTP buffer. For further experiments, the protein solution was desalted using PD10 columns according to the manufacturer's instructions (Amersham Pharmacia Biotech).

Protein concentrations were determined by the dye-binding method of Bradford (1976) using bovine serum albumin as a standard.

**Gel electrophoresis.** PAGE was performed using the Laemmli buffer system (Laemmli, 1970). The purity and the molecular mass of the subunits of GAPDH were examined under denaturing conditions with SDS-polyacrylamide gels. For calibration, standard proteins in the range 14–98 kDa (Silver Stain SDS-PAGE Standards, Low Range; Bio-Rad) were used. Non-denaturing linear gradient gels (4–27.5%) were used to estimate the molecular mass of the native enzyme. Molecular mass standards in the range 67–660 kDa (HMW Calibration kit; Amersham Pharmacia Biotech) were used. Proteins in gels were silver-stained by the method of Blum *et al.* (1987).

**Nucleic acid isolation and manipulation.** Isolation of *C. acetobutylicum* and *C. pasteurianum* chromosomal DNA was performed according to Bertram & Dürre (1989). Plasmid isolation from *E. coli* was done with the Qiagen Midi kit (Qiagen). DNA was manipulated by standard procedures (Sambrook *et al.*, 1989). Enzymes were purchased from commercial suppliers and reactions were carried out under the recommended conditions.

**PCR.** PCR amplifications were performed as described earlier (Treuner-Lange *et al.*, 1997). Synthetic oligonucleotides for PCR and sequencing reactions were prepared by a Gene Assembler Plus (Amersham Pharmacia Biotech) or purchased from commercial suppliers. The following primers were used for PCR amplification of a *gap* gene fragment of *C. pasteurianum*: 5'-GCTATTAATGGTTTTGGAAGAAT-3' (sense primer) and 5'-TGTGATCCATCAACATCTACAAT-3' (antisense primer). These oligonucleotides correspond to the DNA sequence (GenBank accession number X72219) at positions 163–185 and 1061–1039.

**Genomic libraries of *C. acetobutylicum*.** A HindIII genomic library of strain DSM 792 (Sauer *et al.*, 1994) and another genomic library which resulted from partial digestion with *Sau3A*I and ligation into pEcoR252 (Fischer *et al.*, 1993) were used. Colony hybridization using Hybond-N nylon filters (Amersham) was performed for screening the genomic libraries. The filters were prepared according to Buluwela *et al.* (1989) for the hybridization procedure.

**Hybridization.** Chromosomal DNA of *C. acetobutylicum* was digested to completion with the appropriate restriction enzymes, separated on 0.8% agarose gels and transferred to Hybond-N+ nylon membranes (Amersham) by capillary transfer in 20 × SSC (1 × SSC is 0.15 M NaCl, 15 mM sodium citrate). Prehybridization, probe labelling, hybridization, signal generation and detection were performed according to the instructions of the ECL direct nucleic acid labelling and detection system (Amersham).

Colony hybridization was performed with radiolabelled ([ $\alpha$ -<sup>32</sup>P]dATP; Hartmann Analytic) DNA probes. The labelling reaction was performed using the random primer labelling kit (Gibco-BRL Life Technologies), followed by purification of the probe on Sephadex G-25 columns (NAP; Amersham

Pharmacia Biotech). The filters were prehybridized in 0.2% (w/v) polyvinylpyrrolidone, 0.2% (w/v) Ficoll 400, 0.2% (w/v) BSA, 0.1% (w/v) sodium pyrophosphate, 10% (w/v) dextran sulfate, 1 M NaCl, 1% (w/v) SDS and 50 mM Tris/HCl (pH 7.5) for at least 0.5 h at 55 °C. After addition of the denatured, radiolabelled probe, hybridization was continued for 8–12 h. Filters were then washed twice in  $2 \times$  SSC and subjected to autoradiography.

**DNA sequencing and sequence analysis.** A SequiTherm cycle sequencing kit for LI-COR 4000L (Biozym Diagnostik) and IRD41-labelled primers (MWG Biotech) were used for non-radioactive sequencing. The dideoxy-terminated fragments were separated on 41 cm long, 6% polyacrylamide gels or 66 cm long, 4.3% polyacrylamide gels (thickness 0.25 mm) with a LI-COR model 4000L sequencer (MWG Biotech) as instructed by the manufacturer. Data collection and analysis were performed by the programs BASE IMAGIR Data Collection (version 2.31) and BASE IMAGIR Image Analysis (version 2.3).

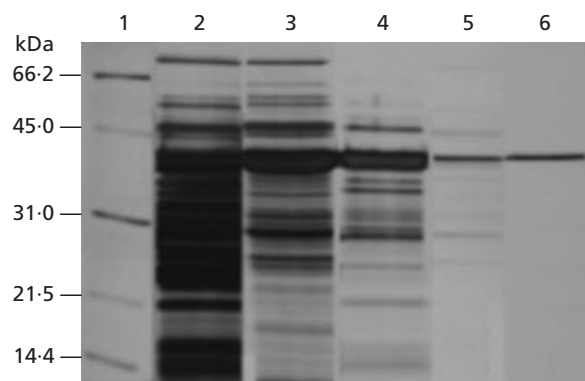
Nucleotide and deduced amino acid sequences were analysed by using the GeneWorks (IntelliGenetics) program on a Power Macintosh 8500/120 computer (Apple). Additional analysis was carried out on a UNIX computer, using the Genetics Computer Group sequence analysis software package (version 9.0).

## RESULTS

### Enzyme purification and characterization

The NAD<sup>+</sup>-dependent GAPDH was purified about 56-fold from acid-producing *C. acetobutylicum* cells using a procedure involving a two-step ammonium sulfate precipitation, gel filtration and anion-exchange chromatography. Results of the purification procedure are summarized in Table 1. The specific activity of the enzyme after anion-exchange chromatography was 27 U mg<sup>-1</sup>. Dye-ligand chromatography on Reactive Blue 72 as last polishing step resulted in a loss of about 80% of the enzyme activity. Characterization was therefore carried out with the desalted anion chromatography pool.

SDS-PAGE experiments revealed one type of subunit with a relative molecular mass of 40 kDa (Fig. 1). The  $M_r$  of native GAPDH was determined by nondenaturing gradient PAGE. Analysis of three separate runs gave an

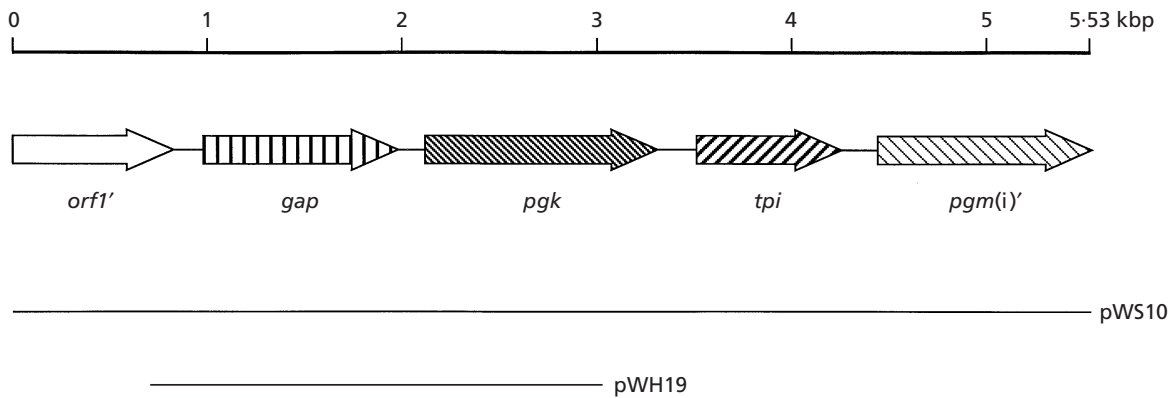


**Fig. 1.** SDS-PAGE of samples from the purification steps of GAPDH from *C. acetobutylicum* DSM 792. The proteins were separated on a 12% acrylamide gel. Lanes: 1, standard proteins (see Methods) with molecular masses on the left; 2, crude extract; 3, ammonium sulfate precipitate; 4, Superdex 200 pool; 5, Q-Sepharose pool; 6, Reactive Blue 72 pool.

average  $M_r$  of 160 000 ( $\pm 10 000$ ; data not shown). In agreement with the structure of many other known NAD<sup>+</sup>-dependent GAPDHs, the enzyme of *C. acetobutylicum* thus shows a homotetrameric composition. A pH activity profile was determined for the physiological direction with glyceraldehyde 3-phosphate (1 mM) as substrate and overlapping buffer systems in the range 5.8–11.6. The polynomial curve showed a maximum at pH 8.5, whereas highest activity was observed with sodium carbonate buffer at pH 9.3. Above pH 10, stability decreased rapidly. The temperature optimum of the enzyme was 50 °C; a further increase resulted in rapid inactivation. Above 65 °C no activity could be detected. According to the Arrhenius plot, an activation energy of 29 kJ mol<sup>-1</sup> was calculated in the range 10–50 °C. GAPDHs of various sources are able to use various aliphatic aldehydes as substrates. Since butyraldehyde is an intermediate in the butanol production of *C. acetobutylicum*, we were interested to determine if the GAPDH is able to use butyraldehyde as substrate. However, in contrast to other GAPDHs (e.g. Fife & Rikihisa, 1970), the clostridial enzyme was not able to

**Table 1.** Purification of GAPDH from *C. acetobutylicum* DSM 792

Purification step	Total protein (mg)	Total activity (U)	Specific activity (U mg <sup>-1</sup> )	Recovery (%)	Purification factor
Crude extract	1500	720	0.48	100	1
Ammonium sulfate (50–90%)	621	518	0.83	72	1.7
Superdex 200	132	495	3.75	69	7.8
Q-Sepharose	6.4	173	27	24	56
Reactive Blue 72	2.5	35	14	5	29



**Fig. 2.** Schematic ORF map of the *gap* gene region in *C. acetobutylicum* DSM 792. The sequenced part of the insert of pWS10 is shown, as well as the insert of pWH19. Arrows represent length, location and orientation of the genes. A prime at the end of an ORF indicates truncation.

oxidize butyraldehyde at a significant rate (less than 1% activity compared to glyceraldehyde 3-phosphate as substrate).

To investigate possible regulation of the enzyme, specific activities were determined in the acidogenic, and early and late solventogenic stages of a continuous culture of *C. acetobutylicum*. The respective values were 0.42, 0.7 and 0.43 U·mg<sup>-1</sup> protein, indicating a low level of regulation at the most.

#### Cloning and sequencing of the *gap* gene region from *C. acetobutylicum*

Based on the close relationship of *C. acetobutylicum* to *C. pasteurianum*, chromosomal DNA of *C. pasteurianum* was used to generate a 0.9 kbp probe via PCR encoding the main part of the *gap* gene. Southern hybridization with this heterologous probe showed positive signals with differently digested chromosomal DNA of *C. acetobutylicum* (data not shown). Two genomic libraries were therefore screened and several positive clones were detected. Two clones were sequenced, one (pWH19) obtained from the *Hind*III genomic library and the other (pWS10) from the partially digested *Sau*3AI genomic library. Fig. 2 shows a schematic map of the *gap* gene cluster. The *gap* gene region of *C. acetobutylicum* includes three complete and two incomplete ORFs, all being transcribed in the same direction. The G + C content of the genes varies between 34 and 35.5 mol%.

From the first ORF (*orf1'*) 831 bp were sequenced. Computer alignments (Table 2) of the deduced amino acid sequence revealed significant homology to the C-terminal part of hypothetical proteins of *Bacillus megaterium* (Urf1), *Bacillus subtilis* (YvbQ) and *Lactobacillus delbrueckii* (YgaP), which are also located upstream of their *gap* genes. The second ORF, encoding a protein of 334 amino acids and a deduced molecular

mass of 35.8 kDa, was located 145 bp downstream of *orf1'*. Because of the high level of similarity of the deduced protein to NAD<sup>+</sup>-dependent GAPDHs of a variety of organisms and especially to the protein of *C. pasteurianum* (95% identity; Table 2), the respective gene was designated *gap*. A putative Shine–Dalgarno sequence (5'-TGGAGG-3') was located 8 bp upstream of the methionine start codon. Separated by 136 bp, a third ORF of 1194 bp was identified downstream of the *gap* gene. The deduced protein comprised 397 amino acids (42.3 kDa). A computer search revealed high level similarity between this gene product and those of various phosphoglycerate kinases (PGKs) (Table 2). The gene, designated *pgk*, is preceded by a putative Shine–Dalgarno sequence (5'-ATGAGG-3') 8 bp upstream of the start codon. Another ORF was located 198 bp downstream of *pgk*. Its predicted amino acid sequence, encoding a protein of 248 amino acids and a molecular mass of 26.5 kDa, showed high similarity to another glycolytic enzyme, triosephosphate isomerase (TPI) (Table 2). The respective gene was therefore designated *tpi*. Eight base pairs upstream of *tpi*, a putative Shine–Dalgarno sequence (5'-AGGAGT-3') was located. Further sequencing revealed a fifth ORF, located 188 bp downstream of *tpi*. Amino acid comparison showed high similarities of the respective gene product with the N-terminal region of members of the subclass of phosphoglycerate mutase (PGM) 2,3-bisphosphoglycerate-independent phosphoglycerate mutase [PGM(i)] (Table 2). A putative Shine–Dalgarno sequence (5'-TGGGGG-3') of this gene, *pgm(i)*, was identified 8 bp upstream of the methionine start codon.

In the intergenic regions no typical rho-independent terminator structures could be identified, suggesting a common transcriptional unit comprising either at least these five genes or rho-dependent termination mechanisms (Platt, 1986). Nucleotide analysis with the programs Stemloop and FoldRNA (Zuker, 1989) revealed only secondary structures with comparatively

**Table 2.** Amino acid identity and similarity of the glycolytic proteins and Orf1' from *C. acetobutylicum* to corresponding proteins of other organisms

Organism	Protein	Identity (%)	Similarity (%)	Reference*
<b>Orf1'</b>				
<i>B. megaterium</i>	Urf1	38	54	1
<i>B. subtilis</i>	YvbQ	38	52	2
<i>L. delbrueckii</i>	YgaP	30	41	3
<b>GAPDH</b>				
<i>C. pasteurianum</i>	GAPDH	95	96	4
<i>S. pyogenes</i>	GAPDH	72	78	5
<i>L. lactis</i> subsp. <i>lactis</i>	GAPDH	68	74	6
<i>E. coli</i>	GAPDH	59	71	7
<i>B. subtilis</i>	GAPDH	56	65	8
<i>S. cerevisiae</i>	GAPDH	47	57	9
<b>PGK</b>				
<i>T. maritima</i>	PGK	62	72	10
<i>B. stearothermophilus</i>	PGK	60	68	11
<i>N. tabacum</i>	PGKY	59	67	12
<i>S. cerevisiae</i>	PGK	53	62	13
<i>C. glutamicum</i>	PGK	48	59	14
<i>E. coli</i>	PGK	46	55	15
<b>TPI</b>				
<i>B. subtilis</i>	TPI	60	67	16
<i>L. lactis</i> subsp. <i>lactis</i>	TPI	59	66	17
<i>B. stearothermophilus</i>	TPI	54	63	18
<i>T. maritima</i>	TPI	52	60	10
<i>C. glutamicum</i>	TPI	46	55	14
<i>E. coli</i>	TPI	45	53	19
<i>S. cerevisiae</i>	TPI	42	51	20
<b>PGM(i)</b>				
<i>B. subtilis</i>	PGM(i)	59	65	16
<i>B. megaterium</i>	PGM(i)	59	65	1
<i>E. coli</i>	PGM(i)	53	62	21

\* 1, Schläpfer & Zuber (1992); 2, Kunst *et al.* (1997); 3, Branny *et al.* (1998); 4, GenBank accession number Q59309; 5, Lottenberg *et al.* (1992); 6, Cancilla *et al.* (1995b); 7, Branlant & Branlant (1985); 8, Viaene & Dhaese (1989); 9, Holland *et al.* (1983); 10, Schurig *et al.* (1995); 11, Davies *et al.* (1991); 12, Rao *et al.* (1995); 13, Hitzeman *et al.* (1982); 14, Alefounder & Perham (1989); 15, Eikmanns (1992); 16, Leyva-Vazquez & Setlow (1994); 17, Cancilla *et al.* (1995a); 18, Rentier-Delrue *et al.* (1993); 19, Pichersky *et al.* (1984); 20, Alber & Kawasaki (1982); 21, Sofia *et al.* (1994).

low energy values. The hairpin with the highest energy value (approx.  $-20.6 \text{ kJ mol}^{-1}$ ) was located downstream of the *gap* gene.

#### Conserved functional and structural domains of the glycolytic genes of *C. acetobutylicum*

The results of alignments of the deduced amino acid sequences of *gap*, *pgk*, *tpi* and *pgm(i)* of *C. acetobutylicum* with the respective gene products of other organisms (Gram-positive and Gram-negative prokaryotes and eukaryotes) are summarized in Table 2. As expected, a comparison of GAPDH of *C. acetobutylicum* revealed high similarity values (95%) to the closely related *C. pasteurianum*. Significant homology

was also observed with GAPDHs of streptococcal origin (Lottenberg *et al.*, 1992). Even enzymes from eukaryotic organisms showed relatively high homology values. Several regions with important catalytic or structural functions are conserved in all sequences, i.e. residues which are involved in forming and stabilizing the nucleotide-binding site (residues 3–18, especially Asn-7, Arg-11, Ile-12, Asn-32, Asp-47, Glu-76, Glu-94, Thr-96, His-108 and Glu-315 in the *C. acetobutylicum* sequence), in forming and function of the binding site for inorganic phosphate (Ser-151, Cys-152 and Thr-153), or in other catalytically important residues such as His-180 (Biesecker *et al.*, 1977; Duée *et al.*, 1996; Moras *et al.*, 1975; Skarzynski *et al.*, 1987; Soukri *et al.*, 1989). A less conserved region is the so-called S-loop, which is

responsible for building the tetrameric structure of GAPDH (residues 182–206).

The mostly monomeric enzyme PGK is also highly conserved among pro- and eukaryotic organisms. The degree of identity of the clostridial enzyme to other PGKs varies between 46% (*E. coli*) and 62% for the PGK part of the PGK–TPI fusion protein of the hyperthermophilic *Thermotoga maritima* (Schurig *et al.*, 1995). In addition, eukaryotic enzymes such as PGKY from *Nicotiana tabacum* show a comparatively high identity of 59%. Taking into account the results of crystal structure and mutant analysis of pro- and eukaryotic organisms, several conserved sequence regions and certain amino acids which are important for the structure and the catalytic mechanism of PGK (Banks *et al.*, 1979; Davies *et al.*, 1993; Fairbrother *et al.*, 1989; Watson *et al.*, 1982) were identified. The PGK of *C. acetobutylicum* also shows the highly conserved residues His-61 and Arg-154 and, in the C-terminal region, amino acids 373–378, which are involved in substrate binding. Arg-64, which is conserved in most sequences, is replaced in the clostridial enzyme by another basic amino acid, lysine. ADP-binding regions could be identified in the *C. acetobutylicum* gene product (residues 198–202, 220–224, 297–298, 319–327 and 352–355, especially Lys-205, Leu-297, Glu-327 and Asp-355). His-61, Arg-154 and His-156 are the counterpart residues in *C. acetobutylicum* which are essential for conformational change of PGK during catalysis (Fairbrother *et al.*, 1989).

TPI is also well characterized by various crystal structure and mutant analyses. The degree of identity of the clostridial TPI and other TPIs varies from 60% (*B. subtilis*) to 42% for the eukaryotic yeast enzyme (Table 2). The catalytically and structurally important residues Lys-11, His-94, Glu-164–Gly 174 (especially Glu-166), Ile-207, residues 209–212 and Leu-231–Leu-237 are also found in the *C. acetobutylicum* sequence (Banner *et al.*, 1975; Lolis *et al.*, 1990; Wierenga *et al.*, 1992).

Based on sequence comparison the clostridial PGM was identified as PGM(i). The first PGM(i) was isolated from maize only in 1992 (Graña *et al.*, 1992), explaining the lack of structural information about this enzyme type. A monomeric composition is characteristic for this class of PGMs (Fothergill-Gillmore & Watson, 1989).

### Functional complementation of an *E. coli gap* mutant

To determine if the predicted clostridial *gap* gene really encodes the purified clostridial NAD<sup>+</sup>-dependent GAPDH, the expression of this gene in the *E. coli gap* mutant (W3CG) was attempted. This mutant carries a transposon insertion within the *gap* gene so that no enzyme activity can be measured nor a complete protein be produced. It is therefore only able to grow with succinate or malate as carbon source. Mutant cells harbouring the recombinant plasmid pWH19 with the whole clostridial *gap* gene and parts of *orf1'* and *pgk* were able to grow on glucose minimal or LB medium.

Enzyme activities were determined from cell extracts of the mutants and the complemented mutants grown on M63 medium with succinate as carbon source. Activity in *C. acetobutylicum* [0.5 U·(mg protein)<sup>-1</sup>] was determined from glucose-grown cultures. Whereas the extracts of W3CG and W3CG harbouring pUC9 showed a specific activity of less than 1 mU·(mg protein)<sup>-1</sup>, the complemented mutant had a 100-fold higher activity [110 mU·(mg protein)<sup>-1</sup>]. The data confirmed that pWH19 really contained the clostridial *gap* gene and that this gene is expressed in the *E. coli* mutant W3CG.

### DISCUSSION

The data reported describe the purification and enzymic characterization of the NAD<sup>+</sup>-dependent GAPDH of *C. acetobutylicum* DSM 792, and cloning and sequencing of the respective gene. Complementation of an *E. coli gap* mutant, and the similarity of the gene product to homologous enzymes, clearly revealed the identity of the gene.

Interestingly, the clostridial GAPDH seems to be a moderately thermotolerant enzyme with a temperature optimum of 50 °C. This could be explained by a slightly higher portion of structure-stabilizing amino acids such as alanine, arginine, leucine and proline (Rentier-Delrue *et al.*, 1993; Schläpfer & Zuber, 1992). An increased thermostability is not characteristic for enzymes of the mesophilic *C. acetobutylicum*. GAPDH does not appear to be significantly regulated during the transition from the acidogenic to the solventogenic stage of *C. acetobutylicum*. The less than twofold increase in specific activity corresponds nicely to data obtained from two-dimensional gel electrophoresis analysis, which also show an about twofold higher abundance of the protein in the early solventogenic stage. However, no increase in synthesis rate could be observed (Schaffer & Dürre, unpublished results), indicating that no regulation at the transcriptional level occurred, but rather a decrease in protein turnover or denaturation. This again is in accordance with the transient increase in chaperone synthesis at the onset of solventogenesis in *C. acetobutylicum* (Pich *et al.*, 1990), which also explains the later decrease in the enzymic specific activity to the original level.

The high degree of conservation of the clostridial GAPDH also extends to residues important for NAD<sup>+</sup>/NADH binding. The invariant Asn-313 of *E. coli* is found in *C. acetobutylicum* at position 314. A change of this amino acid in *E. coli* to threonine resulted in a drastic decrease in NAD<sup>+</sup> affinity (Duée *et al.*, 1996). Another possible way to influence the cofactor effect on the enzyme could be to induce a mutation at Gly-188 Pro-189 (conserved in *E. coli* at positions 187 and 188), which determine the specificity of NAD<sup>+</sup> versus NADP<sup>+</sup>. A change of the proline to a serine in *E. coli* indeed allowed the enzyme to use NADP<sup>+</sup> as a coenzyme (Duée *et al.*, 1996).

The GAPDH of *C. acetobutylicum* shows high homology (72–78%, *S. pyogenes*, Table 2) to members of

this enzyme class that are characterized by additional physiological functions, i.e. adhesion to fibronectin, myosin and actin or a function as plasmin receptor on the cell surface or a role in ADP ribosylation (Lottenberg *et al.*, 1992; Pancholi & Fischetti, 1993). Eukaryotic GAPDHs are also involved in several processes such as endocytosis, DNA repair, tRNA export, programmed death of neurons, etc. (Sirover, 1997). This indicates that GAPDH might possess other functions in addition to its role in energy metabolism.

Nothing is known about the function of the ORF preceding the glycolytic genes in several bacteria (Table 2). It was first detected in *B. megaterium* (Schlöpfer & Zuber, 1992), but there is no unequivocal evidence for its expression. For the corresponding gene in *L. delbrueckii* it was suggested that it is part of an operon localized directly upstream of the *gap-pgk-tpi* operon (Branny *et al.*, 1998). Codon usage analysis of Orf1' shows a much weaker bias compared to the following glycolytic genes, indicating a weaker expression of the putative protein in *C. acetobutylicum*.

Sequence analysis of the incomplete ORF located at the 3'-terminus of the insert revealed high homology to PGM(i)s. Whilst most PGMs show a clear cofactor dependence on bisphosphoglycerate and a monomeric, dimeric or trimeric structure with subunit molecular masses of 23–30 kDa, this newly identified subclass of PGM is characterized by a monomeric structure with a molecular mass of approximately 60 kDa and these enzymes do not show a dependence on bisphosphoglycerate (Fothergill-Gillmore & Watson, 1989). Interestingly, the well-characterized PGM(i) of *B. subtilis* has an absolute requirement for  $Mn^{2+}$ , which is responsible for the extreme sensitivity of the enzyme to pH variations: lowering of pH values results in the inactivation of PGM (Kuhn *et al.*, 1993). This feature of PGM enables the forespore to accumulate 3-phosphoglycerate as an energy storage compound since sporulation is accompanied by a slow decrease in pH. As realkalinization is one of the first reactions during germination, PGM is partly reactivated and 3-phosphoglycerate can be reutilized within the very first minutes of germination (Kuhn *et al.*, 1993). Clostridia also accumulate 3-phosphoglycerate in late phases of sporulation (Loshon & Setlow, 1994) and it is likely that the PGM(i) of *C. acetobutylicum* is also  $Mn^{2+}$ -dependent and plays an important role in accumulating energy resources for germination. It remains to be shown if a decrease in PGM activity in *C. acetobutylicum* takes place upon induction of sporulation or solvent formation, which are accompanied by a decrease in pH.

Organization of glycolytic pathway genes in bacteria is characterized by a high degree of variability. Whereas in some bacteria, for example *L. lactis* (Cancilla *et al.*, 1995a, b), genes like *gap* or *tpi* are not clustered with those of other enzymes of the glycolytic pathway, others show one or more glycolytic genes in the adjacent regions. In *E. coli*, *gapA* and *tpi* are not clustered with other genes of the pathway (Branlant & Branlant, 1985;

Berlyn, 1998), but *pgk*, *fda* (fructose 1,6-bisphosphate aldolase gene) and *gapB* (erythrose-4-phosphate dehydrogenase) are (Alefounder & Perham, 1989). A bicistronic *gap-pgk* operon exists in *Zymomonas mobilis* (Eddy *et al.*, 1989), whereas in *Corynebacterium glutamicum*, *Borrelia burgdorferi* and *Lactobacillus delbrueckii* a cluster of *gap-pgk-tpi* was detected (Eikmanns, 1992; Gebbia *et al.*, 1997; Branny *et al.*, 1998). Exactly the same arrangement of glycolytic genes as in *C. acetobutylicum*, including *orf1'*, was found in *B. subtilis* and *B. megaterium* (Kunst *et al.*, 1997; Schlöpfer & Zuber, 1992). The glycolytic gene cluster of *B. subtilis* contains additionally the gene for the enolase (*eno*) downstream of *pgm(i)*. The possibility that the clostridial gene cluster also contains an additional glycolytic gene downstream of *pgm(i)* cannot be ruled out. Computer analysis revealed no rho-independent terminator structures in this region. Future experiments will aim at a detailed transcript analysis of this region in *C. acetobutylicum*.

## ACKNOWLEDGEMENTS

We thank A. Plückthun for providing the *E. coli gap* mutant W3CG and B. Zickner for technical assistance. This work was supported by a grant from the Deutsche Forschungsgemeinschaft.

## REFERENCES

- Alber, T. & Kawasaki, G. (1982).** Nucleotide sequence of the triose phosphate isomerase gene of *Saccharomyces cerevisiae*. *J Mol Appl Genet* **1**, 419–434.
- Alefounder, P. R. & Perham, R. N. (1989).** Identification, molecular cloning, and sequence analysis of a gene cluster encoding the class II fructose 1,6-bisphosphate aldolase, 3-phosphoglycerate kinase and a putative second glyceraldehyde 3-phosphate dehydrogenase of *Escherichia coli*. *Mol Microbiol* **3**, 723–732.
- Bahl, H., Andersch, W. & Gottschalk, G. (1982).** Continuous production of acetone and butanol by *Clostridium acetobutylicum* in a two-stage phosphate limited chemostat. *Eur J Appl Microbiol Biotechnol* **15**, 201–205 (Erratum **17**, 73, 1983).
- Banks, R. D., Blake, C. C. F., Evans, P. R., Haser, R., Rice, D. W., Hardy, G. W., Merrett, M. & Phillips, A. W. (1979).** Sequence, structure and activity of phosphoglycerate kinase: a possible hinge-bending enzyme. *Nature* **279**, 773–777.
- Banner, D. W., Blommer, A. C., Petsko, G. A. & 9 other authors (1975).** Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution using amino acid sequence data. *Nature* **255**, 609–617.
- Belouski, E., Watson, D. E. & Bennett, G. N. (1998).** Cloning, sequence, and expression of the phosphofructokinase gene of *Clostridium acetobutylicum* ATCC 824 in *Escherichia coli*. *Curr Microbiol* **37**, 17–22.
- Berlyn, M. K. B. (1998).** Linkage map of *Escherichia coli* K-12, edition 10: the traditional map. *Microbiol Mol Biol Rev* **62**, 814–984.
- Bertram, J. & Dürre, P. (1989).** Conjugal transfer and expression of streptococcal transposons in *Clostridium acetobutylicum*. *Arch Microbiol* **151**, 551–557.
- Biesecker, G., Harris, J. I., Thierry, J. C., Walker, J. E. & Wonacott, A. J. (1977).** Sequence and structure of glyceraldehyde-3-phos-

- phate dehydrogenase from *Bacillus stearothermophilus*. *Nature* **266**, 328–333.
- Blum, H., Beier, H. & Gross, H. J. (1987).** Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* **8**, 93–99.
- 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.
- Branlant, G. & Branlant, C. (1985).** Nucleotide sequence of the *Escherichia coli* *gap* gene. Different evolutionary behavior of the NAD<sup>+</sup>-binding domain and of the catalytic domain of D-glyceraldehyde-3-phosphate dehydrogenase. *Eur J Biochem* **150**, 61–66.
- Branny, P., de la Torre, F. & Garel, J.-R. (1998).** An operon encoding three glycolytic enzymes in *Lactobacillus delbrueckii* subsp. *bulgaricus*: glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and triosephosphate isomerase. *Microbiology* **144**, 905–914.
- Buluwela, L., Förster, A., Boehm, T. & Rabbits, T. H. (1989).** A rapid procedure for colony screening using nylon filters. *Nucleic Acids Res* **17**, 452.
- Cancilla, M. R., Davidson, B. E., Hillier, A. J., Nguyen, N. Y. & Thompson, J. (1995a).** The *Lactococcus lactis* triosephosphate isomerase gene, *tpi*, is monocistronic. *Microbiology* **141**, 229–238.
- Cancilla, M. R., Hillier, A. J. & Davidson, B. E. (1995b).** *Lactococcus lactis* glyceraldehyde-3-phosphate dehydrogenase gene, *gap*: further evidence for strongly biased codon usage in glycolytic pathway genes. *Microbiology* **141**, 1027–1036.
- Davies, G. J., Littlechild, J. A., Watson, H. C. & Hall, L. (1991).** Sequence and expression of the gene encoding 3-phosphoglycerate kinase from *Bacillus stearothermophilus*. *Gene* **109**, 39–45.
- Davies, G. J., Gamblin, S. J., Littlechild, J. A. & Watson, H. C. (1993).** The structure of a thermally stable 3-phosphoglycerate kinase and a comparison with its mesophilic equivalents. *Proteins* **15**, 283–289.
- Duée, E., Olivier-Deyris, L., Fanchon, E., Corbier, C., Branlant, G. & Dideberg, O. (1996).** Comparison of the structure of wild-type and a N313T mutant of *Escherichia coli* glyceraldehyde-3-phosphate dehydrogenase: implication for NAD binding and cooperativity. *J Mol Biol* **25**, 814–838.
- Dürre, P. (1998).** New insights and novel developments in clostridial acetone/butanol/isopropanol fermentation. *Appl Microbiol Biotechnol* **49**, 639–648.
- Eddy, C. K., Mejia, J. P., Conway, T. & Ingram, L. O. (1989).** Differential expression of *gap* and *pgk* genes within the *gap* operon of *Zymomonas mobilis*. *J Bacteriol* **171**, 6549–6554.
- Eikmanns, B. J. (1992).** Identification, sequence analysis, and expression of a *Corynebacterium glutamicum* gene cluster encoding the three glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, and triosephosphate isomerase. *J Bacteriol* **174**, 6076–6086.
- Fairbrother, W. J., Walker, P. A., Minard, P., Littlechild, J. A., Watson, H. C. & Williams, R. J. P. (1989).** NMR analysis of site-specific mutants of yeast phosphoglycerate kinase. *Eur J Biochem* **183**, 57–67.
- Fife, T. H. & Rikihisa, T. (1970).** Reaction of glyceraldehyde 3-phosphate dehydrogenase with aliphatic aldehydes. *Biochemistry* **9**, 4064–4067.
- Fischer, R.-J., Helms, J. & Dürre, P. (1993).** Cloning, sequencing, and molecular analysis of the *sol* operon of *Clostridium acetobutylicum*, a chromosomal locus involved in solventogenesis. *J Bacteriol* **175**, 6959–6969.
- Fothergill-Gillmore, L. A. & Watson, H. C. (1989).** The phosphoglycerate mutases. *Adv Enzymol* **62**, 227–313.
- Ganter, C. & Plückthun, A. (1990).** Glycine to alanine substitutions in helices of glyceraldehyde-3-phosphate dehydrogenase: effects on stability. *Biochemistry* **29**, 9395–9402.
- Gebbia, J. A., Backenson, P. B., Coleman, J. L., Anda, P. & Benach, J. L. (1997).** Glycolytic enzyme operon of *Borrelia burgdorferi*: characterization and evolutionary implications. *Gene* **188**, 221–228.
- Girbal, L. & Soucaille, P. (1994).** Regulation of *Clostridium acetobutylicum* metabolism as revealed by mixed-substrate steady-state continuous culture: role of NADH/NAD ratio and ATP pool. *J Bacteriol* **176**, 6433–6438.
- Graña, X., de Lecea, L., El-Maghrabi, R., Ureña, J., Caellas, C., Carreras, J., Puigdomènech, P., Pilikis, S. & Climent, F. (1992).** Cloning and sequencing of cDNA encoding 2,3-bisphosphoglycerate-independent phosphoglycerate mutase from maize. *J Biol Chem* **267**, 12797–12803.
- Hitzeman, R. A., Hagie, F. E., Hayflick, J. S., Chen, C. Y., Seeburg, P. H. & Derynck, R. (1982).** The primary structure of *Saccharomyces cerevisiae* gene for 3-phosphoglycerate kinase. *Nucleic Acids Res* **10**, 7791–7808.
- Holland, J. P., Labieniec, L., Swimmer, C. & Holland, M. J. (1983).** Homologous nucleotide sequences at the 5' termini of messenger RNAs synthesized from the yeast enolase and glyceraldehyde-3-phosphate dehydrogenase gene families. The primary structure of a third yeast glyceraldehyde-3-phosphate dehydrogenase gene. *J Biol Chem* **258**, 5291–5299.
- Jones, D. T. & Woods, D. R. (1986).** Acetone-butanol fermentation revisited. *Microbiol Rev* **50**, 484–524.
- Kuhn, N. J., Setlow, B. & Setlow, P. (1993).** Manganese(II) activation of 3-phosphoglycerate mutase of *Bacillus megaterium*: pH-sensitive interconversion of active and inactive forms. *Arch Biochem Biophys* **306**, 342–349.
- Kunst, F., Ogasawara, N., Moszer, I. & 148 other authors (1997).** The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249–256.
- Laemmli, U. K. (1970).** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Leyva-Vazquez, M. A. & Setlow, P. (1994).** Cloning and nucleotide sequences of the genes encoding triose phosphate isomerase, phosphoglycerate mutase, and enolase from *Bacillus subtilis*. *J Bacteriol* **176**, 3903–3910.
- Lolis, E., Alber, T., Davenport, R. C., Rose, D., Hartman, F. C. & Petsko, G. A. (1990).** Structure of yeast triosephosphate isomerase at 1.9-Å resolution. *Biochemistry* **29**, 6609–6618.
- Loshon, C. A. & Setlow, P. (1994).** Levels of small molecules in dormant spores of *Sporosarcina* species and comparison with levels in spores of *Bacillus* and *Clostridium* species. *Can J Microbiol* **39**, 259–262.
- Lottenberg, R., Broder, C. C., Boyle, M. D., Kain, S. J., Schroeder, B. L. & Curtiss, R. (1992).** Cloning, sequence analysis, and expression in *Escherichia coli* of a streptococcal plasmin receptor. *J Bacteriol* **174**, 5204–5210.
- Lovitt, R. W., Chen, G. J. & Zeikus, J. G. (1988).** Ethanol production by thermophilic bacteria: biochemical basis for ethanol and hydrogen tolerance in *Clostridium thermohydrosulfuricum*. *J Bacteriol* **170**, 2809–2815.
- Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C. & Rossmann, M. G. (1975).** Studies of asymmetry in the three-dimensional structure of lobster D-glyceraldehyde-3-phosphate dehydrogenase. *J Biol Chem* **250**, 9137–9162.

- O'Brien, R. W. & Morris, J. G. (1971). Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. *J Gen Microbiol* **68**, 307–318.
- Pancholi, V. & Fischetti, V. A. (1993). Glyceraldehyde-3-phosphate dehydrogenase on the surface of group A streptococci is also an ADP-ribosylating enzyme. *Proc Natl Acad Sci USA* **90**, 8154–8158.
- Pich, A., Narberhaus, F. & Bahl, H. (1990). Induction of heat shock proteins during initiation of solvent formation in *Clostridium acetobutylicum*. *Appl Microbiol Biotechnol* **33**, 697–704.
- Pichersky, E., Gottlieb, L. D. & Hess, J. F. (1984). Nucleotide sequence of the triosephosphate isomerase gene of *Escherichia coli*. *Mol Gen Genet* **195**, 314–320.
- Platt, T. (1986). Transcription termination and the regulation of gene expression. *Annu Rev Biochem* **55**, 339–372.
- Rao, S. K., Bringloe, D. H., Dyer, T. A., Raines, C. A. & Bradbeer, J. W. (1995). Nucleotide sequences of cDNAs encoding the chloroplastic and cytosolic phosphoglycerate kinases (accession nos. Z48976 and Z48977) from tobacco. *Plant Physiol* **109**, 1126.
- Rentier-Delrue, F., Mande, S. C., Moyens, S., Terpstra, P., Mainfroid, V., Goraj, K., Lion, M., Hol, W. G. & Martial, J. A. (1993). Cloning and overexpression of the triosephosphate isomerase genes from psychrophilic and thermophilic bacteria. Structural comparison of the predicted protein sequences. *J Mol Biol* **229**, 85–93.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sauer, U., Treuner, A., Buchholz, M., Santangelo, J. D. & Dürre, P. (1994). Sporulation and primary sigma factor homologous genes in *Clostridium acetobutylicum*. *J Bacteriol* **176**, 6572–6582.
- Schläpfer, B. S. & Zuber, H. (1992). Cloning and sequencing of the genes encoding glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and triosephosphate isomerase (*gap* operon) from the mesophilic *Bacillus megaterium*: comparison with corresponding sequences from thermophilic *Bacillus stearothermophilus*. *Gene* **122**, 53–62.
- Schurig, H., Beaucamp, N., Ostendorp, R., Jaenicke, R., Adler, E. & Knowles, J. R. (1995). Phosphoglycerate kinase and triosephosphate isomerase from the hyperthermophilic bacterium *Thermotoga maritima* form a covalent bifunctional enzyme complex. *EMBO J* **14**, 442–451.
- Silhavy, T. J., German, M. L. & Enquist, L. W. (1984). *Experiments with Gene Fusions*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sirover, M. A. (1997). Role of glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in normal cell function and in cell pathology. *J Cell Biochem* **66**, 133–140.
- Skarzynski, T., Moody, P. C. E. & Wonacott, A. J. (1987). Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 Å resolution. *J Mol Biol* **193**, 171–187.
- Sofia, H. J., Burland, V., Daniels, D. L., Plunkett, G. & Blattner, F. R. (1994). Analysis of the *Escherichia coli* genome. V. DNA sequence of the region from 76.0 to 81.5 minutes. *Nucleic Acids Res* **22**, 2576–2586.
- Soukri, A., Mougin, A., Corbier, C., Wonacott, A., Branlant, C. & Branlant, G. (1989). Role of the histidine 176 residue in glyceraldehyde-3-phosphate dehydrogenase as probed by site-directed mutagenesis. *Biochemistry* **28**, 2586–2592.
- Treuner-Lange, A., Kuhn, A. & Dürre, P. (1997). The *kdp* system of *Clostridium acetobutylicum*: cloning, sequencing, and transcriptional regulation in response to potassium concentration. *J Bacteriol* **179**, 4501–4512.
- Viaene, A. & Dhaese, P. (1989). Sequence of the glyceraldehyde-3-phosphate dehydrogenase gene from *Bacillus subtilis*. *Nucleic Acids Res* **17**, 1251.
- Watson, H. C., Walker, N. P. C., Shaw, P. J. & 9 other authors (1982). Sequence and structure of yeast phosphoglycerate kinase. *EMBO J* **1**, 1635–1640.
- Wierenga, R. K., Noble, M. E. M. & Davenport, R. C. (1992). Comparison of the refined crystal structures of liganded and unliganded chicken, yeast and trypanosomal triosephosphate isomerase. *J Mol Biol* **224**, 1115–1126.
- Zuker, M. (1989). Computer prediction of RNA structure. *Methods Enzymol* **180**, 262–288.

Received 23 December 1998; revised 10 March 1999; accepted 9 April 1999.