The *Lactococcus lactis* triosephosphate isomerase gene, *tpi*, is monocistronic

Michael R. Cancilla,1 Barrie E. Davidson,1 Alan J. Hillier,2 Nga Y. Nguyen3 and John Thompson4

Author for correspondence: Barrie E. Davidson. Tel: +61 3 344 5912. Fax: +61 3 347 7730. e-mail: davidson@biochemistry.unimelb.edu.au

Triosephosphate isomerase (EC 5.3.1.1) from *Lactococcus lactis* was purified to electrophoretic homogeneity. Approximately 3 mg purified enzyme (specific activity 3300 U mg⁻¹) was obtained from 70 g (wet wt) cells. In solution, triosephosphate isomerase (pl 4.0–4.4) was observed to exist as a homodimer (Mᵦ 57000) of noncovalently linked subunits. The sequence of the first 37 amino acid residues from the NH₂-terminus were determined by step-wise Edman degradation. This sequence, and that of a region conserved in all known bacterial triosephosphate isomerases, was used to design oligonucleotide primers for the synthesis of a lactococcal *tpi* probe by PCR. The probe was used to isolate a molecular clone of *tpi* from a RGEMII library of *L. lactis* LM0230 DNA. The nucleotide sequence of *tpi* predicted a protein of 252 amino acids with the same NH₂-terminal sequence as that determined for the purified enzyme and a subunit Mᵦ of 26802 after removal of the NH₂-terminal methionine. *Escherichia coli* cells harbouring a plasmid containing *tpi* had 15-fold higher triosephosphate isomerase activity than isogenic plasmid-free cells, confirming the identity of the cloned gene. Northern analysis of *L. lactis* LM0230 RNA showed that a 900 base transcript hybridized with *tpi*. The 5' end of the transcript was determined by primer extension analysis to be a G located 64 bp upstream from the *tpi* start codon. These transcript analyses indicated that in *L. lactis*, *tpi* is expressed on a monocistronic transcript. Nucleotide sequencing indicated that the DNA adjacent to *tpi* did not encode another Embden–Meyerhoff–Parnas pathway enzyme. The location of *tpi* on the *L. lactis* DL11 chromosome map was determined to be between map coordinates 1818 and 1978.

**Keywords:** *Lactococcus lactis*, triosephosphate isomerase, transcript analysis, biased codon usage, enzyme purification

INTRODUCTION

The Gram-positive bacterium *Lactococcus lactis* plays an essential role in the manufacture of fermented milk products, such as cheese. The fermentation process requires the uptake of lactose as lactose 6-phosphate by the phosphoenolpyruvate: lactose phosphotransferase system, and the subsequent hydrolysis of lactose 6-phosphate into glucose and galactose 6-phosphate by 6-phospho-β-

Abbreviations: EMP, Embden–Meyerhoff–Parnas; PTH, phenylthiohydantoin.

The GenBank accession number for the nucleotide sequence reported in this paper is U07640.
and conversion of galactose 6-phosphate into triosephosphates have been cloned and characterized (Crow et al., 1983; de Vos et al., 1990; van Rooijen & de Vos, 1990; van Rooijen et al., 1991, 1992). The genes required for the uptake and catabolism of glucose into lactic acid are chromosomally encoded, since plasmid-cured strains of L. lactis can utilize glucose as the sole carbon source (Efstathiou & McKay, 1977). We recently discovered that the L. lactis genes encoding phosphofructokinase (EC 2.7.1.11), pyruvate kinase (EC 2.7.1.40), and lactate dehydrogenase (EC 1.1.1.27) are organized in a single tricistronic transcriptional unit, the las operon (Llanos et al., 1993). This observation is of particular interest since these enzymes are closely regulated at the metabolic level and their relative activities have a major effect on the rates of energy production and lactic acid synthesis in lactic acid bacteria. An operon of this composition has not been reported for any other bacterium and its discovery raises questions about the organization of genes encoding other glycolytic enzymes in L. lactis.

Triosephosphate isomerase is one of the most abundant enzymes in bacterial cells (Noltmann, 1972) and has been described as the near-perfect catalyst (Knowles, 1991). In the eu-bacteria Escherichia coli, Corynebacterium glutamicum and Bacillus megaterium the gene (tpi) encoding the enzyme is located in a cluster of two or more genes encoding glycolytic enzymes (Pichersky et al., 1984; Eikmanns, 1992; Schläpfer & Zuber, 1992). Northern analyses have revealed that this cluster constitutes a polycistronic operon in C. glutamicum (Schwinde et al., 1993).

In view of the role of triosephosphate isomerase in the homofermentative process in L. lactis, we have carried out an analysis of this enzyme and its structural gene. In this report we describe the purification of triosephosphate isomerase from L. lactis to electrophoretic homogeneity, its NH₂-terminal amino acid sequence and some properties of the purified enzyme. The availability of the sequence enabled us to clone the gene encoding triosephosphate isomerase, which we have designated tpi, and to analyse its structure and transcriptional organization.

METHODS

Bacterial strains, plasmids, bacteriophage, and culture conditions. The bacterial strains, plasmids, and bacteriophage used in this study are listed in Table 1. L. lactis strains LM0230 and DL11 were grown at 30 °C in M17G [M17 medium (Terzaghi & Sandine, 1975), without lactose but supplemented with 0.5% (w/v) glucose]. L. lactis K1-42 was grown for purification of triosephosphate isomerase in a previously described medium (Thompson & Saier, 1981). D-Ribose (0.5%, w/v) was used as the carbon source so that the cell extract could also be used to purify fructokinase II, which is derepressed by ribose (Thompson et al., 1991b); triosephosphate isomerase levels are not altered by growth in D-ribose. After overnight growth at 32 °C to stationary phase, the cells were harvested by centrifugation, washed by resuspension and centrifugation from KP buffer (KP buffer is 10 mM potassium phosphate, pH 7.4), and stored at -20 °C until required. The cell yield was approxi-

mately 3-5 g (wet wt) of cells per litre. E. coli was grown in 2YT medium (Miller, 1972) at 37 °C with aeration. Antibiotics were used at the following concentrations: erythromycin, 200–400 µg ml⁻¹; ampicillin, 100 µg ml⁻¹; kanamycin, 25 µg ml⁻¹; and tetracycline, 12.5 µg ml⁻¹.

Chemicals and reagents. Protein standards for SDS-PAGE and gel filtration chromatography, isoelectric focusing standards, Ampholine polyacrylamide gel plates, DEAE-Sephadex and phenyl-Sepharose CL-4B were obtained from Pharmacia. Ultragel AcA 44 was purchased from Sepacor and hydroxyapatite (Fast Flow) from Calbiochem. Dl-Glyceraldehyde 3-phosphate diethyl acetal (monobarium salt), α-glycerophosphate dehydrogenase (type III, rabbit muscle), and other reagents, were purchased from Sigma.

Purification of triosephosphate isomerase

The SDS-PAGE procedure of Laemmli (1970) was used to monitor enzyme purity throughout the purification. Preparation of SDS-polyacrylamide gels, electrophoresis conditions, and staining procedures have been described previously (Thompson, 1989). Protein concentrations were determined by use of the Bio-Rad protein dye-binding reagent with BSA as the standard. Except where indicated, all purification steps were carried out at 4 °C.

(i) Cell disruption and preparation of high-speed supernatant. Frozen cells of L. lactis K1-42 (70 g wet wt) were thawed and resuspended in KP buffer to a volume of 150 ml. The cells were disrupted at 0 °C by three 5 min periods of sonic oscillation with a Branson model 350 sonifier operating at 75% maximum power. Intact cells and debris were removed by centrifugation (25 000 g, 30 min), and the supernatant was dialysed overnight against 4 l KP buffer. The dialysate was concentrated twofold by ultrafiltration through a Diaflo PM-10 membrane (Amicon) and centrifuged (100 000 g, 2 h).

(ii) Phosphocellulose chromatography. The high-speed supernatant (54 ml) was transferred to a 2.6 x 24 cm column of phosphocellulose (Whatman P-11) previously equilibrated with KP buffer. The non-adsorbed proteins, which included triosephosphate isomerase, were eluted with KP buffer. The active fractions (usually fractions 45–51) were pooled and concentrated to 20 ml by ultrafiltration.

(iii) DEAE-Sephadex chromatography. The concentrate from step (ii) was transferred to a 2.6 x 24 cm column of DEAE-Sephadex previously equilibrated with 25 mM HEPES, pH 7.5. After washing the column with 150 ml of this buffer, a 600 ml linear NaCl gradient (0-2-0.5 M) in 25 mM HEPES, pH 7.5, was passed through the column and 5 ml fractions were collected. The active fractions (usually fractions 45–51) were pooled and concentrated to 10 ml by ultrafiltration.

(iv) Phenyl-Sepharose chromatography. Solid NaCl was added to the concentrate from step (iii) to a final concentration of 1 M. The sample was then applied to a 1 x 28 cm column of phenyl-Sepharose CL-4B equilibrated with KP buffer containing 1 M NaCl. Unadsorbed proteins were eluted with this buffer, then a 300 ml linear decreasing gradient of NaCl (0.6-0.1 M) in 25 mM HEPES, pH 7.5, was passed through the column and 2 ml fractions were collected. The active fractions (fractions 10–24) were pooled and concentrated to 3 ml by ultrafiltration.

(v) Hydroxyapatite chromatography. The concentrate from step (iv) was desalted and exchanged for KP buffer by passage through a PD-10 gel filtration column previously equilibrated with KP buffer. The desalted preparation was transferred to a 1 x 5 cm column of hydroxyapatite equilibrated with KP buffer. In contrast to the contaminating proteins, the triosephosphate
Triosephosphate isomerase was not adsorbed by this matrix, and the purified enzyme (about 3 mg) was eluted in the wash fractions and concentrated by ultrafiltration.

**M, determinations.** The M, of native triosephosphate isomerase in solution was determined by the gel filtration method of Andrews (1964) using a 1.6 × 94 cm column of Ultrogel AcA44 equilibrated with 50 mM potassium phosphate buffer, pH 7, containing 0.1 M NaCl. Standard proteins used to calibrate the column were: BSA (M, 66000), ovalbumin (45000), carbonic anhydrase (31000) and chymotrypsinogen A (25000). SDS-PAGE was used to estimate the subunit M, of the denatured triosephosphate isomerase. Standard proteins used to calibrate the gel were: phosphorylase b (97000); BSA, ovalbumin, carbonic anhydrase, soybean trypsin inhibitor (21 500) and lysozyme (14 400).

**Isoelectric focusing and pl estimation.** The isoelectric point (pl) of purified triosephosphate isomerase was determined by analytical isoelectric focusing in a Multiphor model 2117 flat-bed electrophoresis unit (Pharmacia). Electrophoresis conditions, and procedures for fixing, staining and destaining of 1 mm thick isoelectric focusing gels were done according to the manufacturer's instructions. The pl calibration standards (pl range 3-10) and precast Ampholine polyacrylamide gel plates (pH range 3-5-9.5) were purchased from Pharmacia.

**Triosephosphate isomerase assay.** Triosephosphate isomerase catalyses the reversible conversion of D-glyceraldehyde 3-phosphate to dihydroxyacetone phosphate. Enzyme activity was determined spectrophotometrically at 340 nm by monitoring the rate of NADH-dependent, α-glycerophosphate-dehydrogenase-mediated conversion of dihydroxyacetone phosphate (produced from glyceraldehyde 3-phosphate) to α-glycerophosphate. This assay would be perturbed by glyceraldehyde-3-phosphate dehydrogenase and NADH oxidase activities in the sample; however, the relative levels of these enzymes in the high-speed supernatant compared with that of triosephosphate isomerase were negligible. One unit of triosephosphate isomerase was defined as the amount of enzyme that catalysed the oxidation of 1 pmol NADH (i.e. equivalent to the conversion of 1 pmol glyceraldehyde 3-phosphate to dihydroxyacetone phosphate) per min at 22 °C. The standard 1 ml assay mixture contained 0.1 M imidazole/HCl buffer (pH 7.0), 0.1 mM NADH, 10 mM MgCl2, 0.25 mM D-glyceraldehyde 3-phosphate, 5 U α-glycerophosphate dehydrogenase (EC 1.1.1.8) and triosephosphate isomerase.

**NH₂-terminal amino acid sequence.** Approximately 100 pmol dry, salt-free protein was dissolved in 30 μl sequencing-grade trifluoroacetic acid (0-1%, v/v), and dried on a glass fibre filter membrane which had been coated with polybrene. The protein was subjected to three filter cycles with an Applied Biosystems model 470A gas phase microsequencer, and PTH-amino acid derivatives were identified on an Applied Biosystems model 120A analyser.

**DNA techniques and sequence analysis.** *E. coli* cells were transformed by the method of Hanahan (1983). Plasmid DNA from *E. coli* was isolated by alkaline-lysis procedure (Birnboim & Doly, 1979) and was purified by the addition of PEG (to

### Table 1. Bacterial strains, plasmids and bacteriophage used

<table>
<thead>
<tr>
<th>Strains</th>
<th>Characteristics</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>L. lactis</em> K1-42</td>
<td>L. lactis subsp. lactis K1 after spontaneous loss of lac plasmid and Tn5306</td>
<td>Donkersloot &amp; Thompson (1990); Thompson et al. (1991a)</td>
</tr>
<tr>
<td>LM0230</td>
<td>Plasmid-cured derivative of <em>L. lactis</em> subsp. cremoris C2 (formerly <em>L. lactis</em> subsp. lactis C2)</td>
<td>Efstathiou &amp; McKay (1977); Salama et al. (1991)</td>
</tr>
<tr>
<td>DL11</td>
<td>Pre+ derivative of <em>L. lactis</em> subsp. lactis ATCC 11454</td>
<td>Tulloch et al. (1991)</td>
</tr>
<tr>
<td><em>E. coli</em> JM107</td>
<td>endA1 gyrA96 thi-1 hsdR17 supE44 relA1 Δ(lac-proAB) (F' traD36 proAB lacYZ ΔM15)</td>
<td>Yanisch-Perron et al. (1985)</td>
</tr>
<tr>
<td>SURE</td>
<td>c14Δ(mcrA) Δ(mcrCB-bsdSMR-mrr)171 ielC recR recJ umuC::Tn5(Km') worC supE44 lac gyrA96 relA1 thi-1 endA1(F' pro AB lacY ΔM15)</td>
<td>Stratagene</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGEM-T</td>
<td>ApΔΔlacZ, 3005 bp vector for cloning PCR products</td>
<td>Promega</td>
</tr>
<tr>
<td>pJDC9</td>
<td>EmΔΔlacZ, 6-85 kbp cloning vector</td>
<td>Chen &amp; Morrison (1988)</td>
</tr>
<tr>
<td>pMU5501</td>
<td>ApΔΔlacZ with 503 bp of tpi prepared by PCR amplification of <em>L. lactis</em> LM0230 DNA</td>
<td>This study</td>
</tr>
<tr>
<td>pMU5502</td>
<td>EmΔ, pGEM-T with a 6-0 kbp EcoRI fragment containing tpi from <em>L. lactis</em> LM0230</td>
<td>This study</td>
</tr>
<tr>
<td>pMU5503</td>
<td>EmΔ, pJDC9 with a 24 kbp Dral fragment containing tpi from <em>L. lactis</em> LM0230</td>
<td>This study</td>
</tr>
<tr>
<td>Bacteriophage λ268</td>
<td>λGEM11 with a 21-0 kbp partial Sau3AI fragment from <em>L. lactis</em> LM0230 in the XbaI site</td>
<td>This study</td>
</tr>
</tbody>
</table>
nucleotides were synthesized in an Applied Biosystems model M. R. CANCILLA

381A DNA synthesizer. Bacteriophage DNA used as template in the PCR was isolated from L. LM0230 as described previously (Cancilla

DNA strands using an Applied Biosystems model 373A

amino acid sequences were analysed using the programs on the University of Sydney SUN computer system. 

Construction of a molecular probe for tpi. The following oligonucleotides were used as polymerase chain reaction (PCR) primers: N-TPI, 5'-GG(GACTG)AA(C/CT)GGAA(A/AG)AT(GT)(CC)GAT- AA-3', and C-TPI, 5'-CC(GACTG)GT(GACTG)CCGAT- (ACTG)GCCC-3'. N-TPI encoded amino acid residues 8–13 (GNWKMN) of L. lactis K1-42 triosephosphate isomerase, and the complementary strand of C-TPI encoded WAIGTG, a conserved sequence in all known bacterial triosephosphate isomerases near residues 171–176. The PCR mix (50 μl final volume) consisted of template DNA (10–20 ng), N-TPI and C-TPI (250 ng each), deoxynucleoside triphosphates (200 μM each), 1 × Taq DNA polymerase buffer (Boehringer Mannheim), and 2 U Taq DNA polymerase (Boehringer Mannheim). The reaction was carried out in a thermal sequencer (model FTS-1C, Corbett Research) as follows: 95 °C for 3 min, and then 30 cycles of 45 °C for 1 min, 72 °C for 1.5 min, and 95 °C for 0.5 min. The amplified product was isolated and cloned into pGEM-T.

Northern analysis. Total RNA was purified from L. lactis LM0230, electrophoresed, transferred to a nylon membrane and hybridized as described previously (Chandy et al., 1994). The size of the tpi transcript was determined by comparison of its electrophoretic mobility with those of a set of RNA markers of Lactococcus lactis, K1-42 triosephosphate isomerase, and other known triosephosphate isomerase amino acid sequences. Primer extension analysis. Total RNA (5 μg) was annealed with 750 ng of the oligonucleotide tpi2 (5'-GCACCGATAAC- TGATTG-3') in 5 μl water, by incubation at 90 °C for 1 min then 42 °C for 5 min. Radiolabelled cDNA was synthesized by using [33P]dATPαS and AMV reverse transcriptase (Boehringer Mannheim), according to the method of Chandy et al. (1994). The products of the four different dideoxy sequencing reactions obtained with the fmol DNA sequencing system (Promega) using pMU5503 as the template and tpi2 as the primer, served as the size standards on the 6% (w/v) polyacrylamide/7 M urea gel.

RESULTS

Purification and properties of triosephosphate isomerase

Cell-free extracts of L. lactis K1-42 grown with ribose exhibited high levels of triosephosphate isomerase activity (60 U mg⁻¹), and were a suitable starting material for purification of the enzyme (Table 2). The five-stage purification procedure usually provided 3–4 mg of electrophoretically pure enzyme from 70 g (wet wt) L. lactis cells. The purified enzyme remained stable for several months when stored at −20 °C in 50 mM potassium phosphate buffer (pH 7.0), exhibiting little loss of activity upon repeated (at least four times) freezing and thawing. The Mₐ of native triosephosphate isomerase was estimated to be 57000 by Ultrogel AcA44 filtration chromatography (data not shown), while SDS-PAGE of the fully reduced enzyme revealed a single polypeptide of Mₐ 27000 (Fig. 1, lane 5). These data indicated that triosephosphate isomerase from L. lactis exists in the native state as a homodimer of noncovalently linked subunits. Although SDS-PAGE of purified triosephosphate isomerase revealed a single protein band, the same preparation was resolved into two species by analytical electrofocusing (data not shown). The pls were estimated to be 4.0 for the more sharply focused polypeptide, and 4.4 for the more diffuse species. Whether this microheterogeneity reflected post-translational modification(s), deamidation or proteolysis during purification, or whether it was an artefact of the electrofocusing procedure, is unclear. Microsequencing of the pure triosephosphate isomerase provided a reasonably unambiguous sequence for the first 37 NH₂-terminal residues: Ser-Arg-Lys-Pro-Ile-Ile-Ala-Gly-Asn-(Trp/Cys)-Lys-Met-Asn-Lys-Thr-Leu-(Ala/Ser)-Glu-Ala-Gln-Ala-Phe-Val-Glu-Ala-Val-Lys-Asn-Asn-Leu-Pro-(Pro/Ser)-Ser-Asp-Asn-Val-Glu.

Molecular cloning and sequence analysis of tpi

The availability of the NH₂-terminal sequence of L. lactis triosephosphate isomerase enabled the use of PCR to generate a specific probe for the L. lactis gene (which we designated tpi) encoding the enzyme. One primer, N-TPI, was designed from the NH₂-terminal sequence data and the other, C-TPI, was designed from the sequence WAIGTG, which is conserved in all known bacterial triosephosphate isomerase sequences. PCR amplification of L. lactis DNA with these primers yielded a fragment of 503 bp, which is the approximate size expected from known triosephosphate isomerase amino acid sequences. The fragment was cloned (pMU5501) (Table 1) and its nucleotide sequence was determined. The predicted amino acid sequence indicated that the fragment contained the expected region of tpi. The 503 bp fragment from pMU5501 was then used as a specific probe to screen a JGEM11 genomic library of L. lactis LM0230 (Llanos et al., 1992). One clone (2668) hybridized strongly with the probe. DNA was isolated from this clone and Southern analyses were used to construct a restriction map of a 60 kb EcoRI fragment located within the 2668 L. lactis DNA insert (Fig. 2). The 60 kb EcoRI fragment that hybridized with the tpi probe was subcloned from 2668 into pJDC9, yielding a recombinant plasmid containing tpi and its flanking DNA (pMU5502).

HaeIIII, DraI, HindIII, SspI and PvuII fragments from the 60 kb insertion in pMU5502 were cloned in pJDC9, and the nucleotide sequences of these fragments were determined. The 1138 bp nucleotide sequence shown in Fig.
Table 2. Purification of triosephosphate isomerase from *Lactococcus lactis* K1-42

<table>
<thead>
<tr>
<th>Purification step*</th>
<th>Enzyme activity (U)</th>
<th>Amount of protein (mg)</th>
<th>Sp. act. (U mg⁻¹)</th>
<th>Purification (-fold)</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-speed supernatant</td>
<td>127 000</td>
<td>2160</td>
<td>60</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Phosphocellulose P-11 chromatography</td>
<td>84 500</td>
<td>1460</td>
<td>60</td>
<td>1</td>
<td>67</td>
</tr>
<tr>
<td>DEAE-Sephacel chromatography</td>
<td>31 100</td>
<td>220</td>
<td>140</td>
<td>2:3</td>
<td>25</td>
</tr>
<tr>
<td>Phenyl-Sepharose chromatography</td>
<td>28 700</td>
<td>12:3</td>
<td>2300</td>
<td>38</td>
<td>23</td>
</tr>
<tr>
<td>Hydroxylapatite chromatography</td>
<td>11 500</td>
<td>3:5</td>
<td>3300</td>
<td>55</td>
<td>9</td>
</tr>
</tbody>
</table>

* See Methods for details.

![Triosephosphate isomerase from *Lactococcus lactis*](image)

**Fig. 1.** Purification and *M*, determination of triosephosphate isomerase from *Lactococcus lactis* K1-42. SDS-PAGE (12.5%, w/v, acrylamide) of samples from each stage of triosephosphate isomerase purification: high-speed supernatant (lane 1); phosphocellulose (lane 2); DEAE-Sephacel (lane 3); phenyl-Sepharose (lane 4); hydroxylapatite (lane 5). The numbers on the right are the molecular sizes of standard proteins (see Methods) electrophoresed in the right and left hand lanes. The arrow indicates the subunit of purified triosephosphate isomerase in lane 5 (*M*, 27000).

![Restriction map](image)

**Fig. 2.** Restriction map of the 60 kbp EcoRI insert in pMU5502 containing *Lactococcus lactis* tpi. The digestion sites shown are for EcoRI (E), HindIII (H), PvuII (P), SspI (Ss), Dral (D), Sau3AI (Sa), and HaellI (Ha). The positions of tpi, ORFs a, b and c are indicated by rectangles, with arrows showing their orientations. Solid portions of the rectangles indicate sequenced regions. The locations of the tpi probe generated by PCR and the oligonucleotide tpi2 used in primer extension analysis are indicated.

Expression of *Lactococcus lactis* tpi in *E. coli*

The predicted amino acid sequence (Fig. 3) provided evidence that the lactococcal DNA in pMU5502 contained tpi. To establish this point, we looked for the expression of the *Lactococcus lactis* tpi gene in *E. coli* cells harbouring pMU5502. Cell extracts from *E. coli* SURE(pMU5502) for triosephosphate isomerase (Fig. 1). The PI of the protein predicted from the tpi sequence was 4.4, which is also in agreement with the values that were determined experimentally. A putative ribosome-binding site was identified 8 bp upstream from the tpi initiation codon (Fig. 3). The 3' flanking DNA of tpi contained an inverted-repeat sequence followed by TTT (Fig. 3). Transcripts of this region could form a stem-loop structure with a calculated free energy value (Tinoco *et al.*, 1973) of -15.6 kcal mol⁻¹ (-65.3 kJ mol⁻¹). A stem-loop structure followed by UUU is likely to be a rho-independent transcription terminator (d'Aubenton Carafa *et al.*, 1990), causing RNA polymerase to stop transcribing at bp 988 (Fig. 3).
had a 15-fold higher level of triosephosphate isomerase (81 U mg⁻¹) than those from plasmid-free *E. coli*. The translated region of 253 codons starts with an ATG at nucleotide positions 178-180. The amount of a polypeptide of the expected size increased amount of a polypeptide of the expected size expressed in *E. lactis* Yo by using the determined specific activity of pure *L. lactis* tpi and its flanking DNA and the deduced amino acid sequence of *L. lactis* tpi2 using total RNA as template yielded one transcript (Fig. 4a). Primer extension from oligo-nucleotide tpi2 using total RNA as template yielded one cDNA product (Fig. 4b). The mobility of the product was identical to that of the product in the C track, corresponding to bp 115 in Fig. 3. We concluded from this observation that the G at bp 115 was the 5' end of the tpi transcript, and indicates that *L. lactis* tpi is expressed as a monocistronic mRNA.

### Sequence analysis of the DNA adjacent to tpi

Subcloning experiments from pMU5502 provided constructs which allowed the determination of partial nucleotide sequences of the DNA adjacent to tpi (Fig. 2). Computer analysis of these sequences was used to identify potential ORFs. The predicted amino acid sequences of these ORFs (designated a, b, and c, Fig. 2) were compared with the sequences in the GenPeptide data base using the FASTA program (Lipman & Pearson, 1985). ORFs a and b shared 30% identity over a combined total of 145 amino acids with enzyme II encoded by ptsG from *Brevibacterium lactoflavorum* (*C. glutamicum*) (GenBank accession no. L18875), and ORF c shared 25% identity over 147 amino acids with the conjugated bile acid hydrolase encoded by cbb from *Lactobacillus plantarum* (GenBank accession no. M96175). Neither ORF exhibited any significant homology with known EMP pathway enzymes. Although additional nucleotide sequence and biochemical analyses need to be undertaken to establish the identity of ORFs a, b, and c, the present data indicate that other genes encoding EMP pathway enzymes are not adjacent to tpi.

### Transcript analyses

Northern analysis of *L. lactis* LM0230 RNA demonstrated that a tpi probe hybridized strongly with a single 900 base transcript (Fig. 4a). Primer extension from oligonucleotide tpi2 using total RNA as template yielded one cDNA product (Fig. 4b). The mobility of the product was identical to that of the product in the C track, corresponding to bp 115 in Fig. 3. We concluded from this observation that the G at bp 115 was the 5' end of the tpi transcript; a transcript extending from this 5' end to its 3' end equivalent at bp 988 would be 873 bases long. This value agrees with the experimentally determined size of 900 bases for the tpi transcript, and indicates that *L. lactis* tpi is expressed as a monocistronic mRNA.

### Codon usage in *L. lactis* tpi

Codon usage in *L. lactis* tpi exhibits a marked bias towards particular codons and away from others. Of the 61 possible amino acid codons, 24 are not used in this gene. While the low G+C content (37 mol%) of *L. lactis* DNA may contribute to the non-usage of codons with a G or C in the third position, it cannot be the reason for the preference for T compared with A in the third position of Thr, Leu, Ile, Ala, Arg, and Gly codons. These data, when combined with compilations of codon usage for other genes encoding glycolytic enzymes (*lac operon*) (Llanos et al., 1993) suggest that a strong bias exists in the codon usage of glycolytic genes in *L. lactis*.
Copy number and chromosomal location of tpi

EcoRI, DraI, HaeIII, NruI, and Sphi digests of L. lactis LM0230 chromosomal DNA were resolved by electrophoresis through agarose, and transferred to a nylon membrane. Hybridization of the membrane with a 32P-labelled tpi probe (data not shown) yielded the following hybridizing fragments: 6-0 kbp (with EcoRI), 1-8 kbp (DraI), 1-0 kbp (HaeIII), > 100 kbp (NruI), and 1-5 and 1-35 kbp (Sphi). The measured sizes of these fragments corresponded with the sizes predicted from restriction mapping (Fig. 2) and sequence data (Fig. 3). These observations indicated the presence of a single copy of tpi in the L. lactis LM0230 chromosome.

A physical map of the L. lactis DL11 chromosome has been constructed by pulsed-field gel electrophoresis (Tulloch et al., 1991). To locate tpi on this map, samples of high-Mr DNA from L. lactis DL11 were digested with SmaI, NatI plus SmaI, and NosI, then separated by pulsed-field gel electrophoresis, and transferred to a nylon membrane. The membrane was probed with 32P-labelled tpi and fragments of 830, 160, and 160 kbp were observed to hybridize (data not shown). The 830 and 160 kbp bands corresponded to fragments SmA and NtF, respectively (Tulloch et al., 1991). Since it has previously been shown that NtF is located within SmA, we concluded that the single copy of tpi is located on the L. lactis DL11 chromosome within fragment NtF, between coordinates 1-818 and 1-978 on the physical map (Tulloch et al., 1991).

DISCUSSION

The data reported above describe the enzymic and genetic characterization of triosephosphate isomerase from L. lactis. Electrophoretically homogeneous enzyme was purified from L. lactis and its NH2-terminal amino acid sequence was determined. PCR was then used to synthesize a specific lactococcal probe that could be used to isolate a molecular clone of tpi, the gene encoding triosephosphate isomerase. The NH2-terminal sequence of the purified enzyme was utilized to design one low-redundancy PCR primer while the sequence WAIGTG, which is conserved in all known triosephosphate isomerases (Fig. 5; Fothergill-Gilmore & Michels, 1993), was used for the second.

Proof that our clone encoded triosephosphate isomerase was obtained by (i) the identity of its predicted NH2-terminal amino acid sequence with that determined for the purified enzyme, and (ii) the observation of elevated levels of triosephosphate isomerase activity in E. coli cells that contained the clone. The purified enzyme was shown to exist in solution as a homodimer (Mr, 57 000) of non-covalently linked subunits, and the sequence analysis of tpi indicated that the enzyme subunits are identical and have an Mr of 26 802. The codon usage in L. lactis tpi was found to exhibit a strong bias. This was in direct agreement with the codon usage for genes of other L. lactis chromosomal genes encoding EMP pathway enzymes (Llanos et al., 1993), and noticeably different to data compiled from 67 other lactococcal genes (Chopin, 1993). By analogy with the situation in E. coli (Gouy & Gautier, 1982) and yeast (Sharp et al., 1986), this observation suggests that tpi is a strongly expressed gene.

The predicted amino acid sequence of L. lactis triosephosphate isomerase shares considerable identity with the amino acid sequences of other bacterial triosephosphate isomerases (Fig. 5). The identity ranges from 60% with the Bacillus megaterium enzyme down to 33% with the enzyme from Moraxella sp. Extensive structural and mutagenic studies with triosephosphate isomerase from various organisms have resulted in the identification of amino acid residues involved in subunit interaction, ligand binding and catalysis (Knowles, 1991; Lolis et al., 1990; Wierenga et al., 1992). Each subunit contains a catalytic centre that consists of a number of highly conserved amino acid residues (Wierenga et al., 1992). These conserved amino acids (Lys-12, His-96 and Glu-168) were readily identified in the predicted sequence of triosephosphate isomerase from L. lactis (Fig. 5).

Northern blot analyses with L. lactis RNA showed that tpi is expressed in a 900 base transcript. Given the length of
the triosephosphate isomerase coding region (756 bp), and the distance between the start and termination sites of transcription (873 bp) we conclude that tpi is monocistronic. In a number of other bacterial species tpi has been found to be adjacent to genes encoding other enzymes of the EMP pathway. For example, in the Gram-positive bacteria C. glutamicum and B. megaterium a gene cluster of gap encoding glyceraldehyde-3-phosphate dehydrogenase, EC 1.2.1.12), pgk (encoding phosphoglycerate kinase, EC 2.7.2.3) and tpi have been discovered (Eikmanns, 1992; Rentier-Delrue et al., 1993). The three enzymes encoded by the cluster catalyse sequential reactions in the pathway. Transcriptional analyses of the gap-pgk-tpi cluster from C. glutamicum (Eikmanns, 1992, 1993), B. megaterium (Rentier-Delrue et al., 1993b), B. megaterium (Bme) (Rentier-Delrue & Zuber, 1992), L. lactis (Lla) (this study), and the conserved amino acids from the (con). The first amino acid for each protein is the NH₂-terminal amino acid. Spaces indicate positions where sequences are identical to that of L. lactis; dots indicate gaps inserted to maximize the alignment; double underlining indicates catalytic residues in the triosephosphate isomerase active site (Wierenga et al., 1992); asterisks show the COOH-terminal extremity of each sequence; the percentages show the identity between each sequence and the L. lactis sequence.

ACKNOWLEDGEMENTS

We are grateful to Sean Moore for assistance with gel preparation and use of the automated DNA sequencing. This study was supported by the Australian Research Council and the Australian Dairy Research and Development Corporation. M. R. C. acknowledges receipt of an Australian Dairy Research and Development Corporation Postgraduate Studentship.

REFERENCES


