JOURNAL OF BACTERIOLOGY, Apr. 1988, p. 1926–1933
0021-9193/88/041926-08$02.00/0

Vol. 170, No. 4

Phosphoglycerate Kinase Gene from *Zymomonas mobilis*: Cloning, Sequencing, and Localization within the *gap* Operon†

T. CONWAY AND L. O. INGRAM*

Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida 32611

Received 2 October 1987/Accepted 22 January 1988

The *Zymomonas mobilis* gene encoding phosphoglycerate kinase (EC 2.7.3.2), *pgk*, has been cloned into *Escherichia coli* and sequenced. It consists of 336 amino acids, including the N-terminal methionine, with a molecular weight of 41,384. This promoterless gene is located 225 base pairs downstream from the *gap* gene and is part of the *gap* operon. Previous studies have shown that the specific activities of glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate kinase do not change coordinately in *Z. mobilis*, although the two enzymes appear to be under the control of a common promoter. The translated amino acid sequence for the *Z. mobilis* phosphoglycerate kinase is less conserved than those of eucaryotic genes. A comparison of known sequences for phosphoglycerate kinase revealed a high degree of conservation of structure with 102 amino acid positions being retained by all. In general, the amino acid positions at the boundaries of β-sheet and α-helical regions and those connecting these regions were more highly conserved than the amino acid positions within regions of secondary structure.

Phosphoglycerate kinase catalyzes the first glycolytic reaction which produces ATP, the transfer of a high-energy phosphoryl group from the carbonyl of bis-phosphoglycerate to ADP. It is the only monomeric enzyme in glycolysis and is one of the most abundant proteins in the cytoplasm of fermenting yeasts and bacteria (29). This enzyme is also one of the proteins which is overproduced in *Saccharomyces cerevisiae* in response to heat shock (27).

Phosphoglycerate kinase has been the subject of intensive investigation as a model for protein structure and catalysis (3, 29, 33) and for the study of sex-linked hereditary defects in humans (5, 15). The three-dimensional structures of the horse (3) and yeast phosphoglycerate kinase (26) have been determined. These enzymes exhibit a remarkable degree of structural conservation. Each enzyme is folded to form two lobes, connected by a hinge region. The lobe formed by the carboxyl end of the protein is primarily involved in the binding of ADP, whereas the lobe formed by the amino end of the protein binds bis-phosphoglycerate (3, 26). Substrate binding and catalysis are associated with a folding together of these two lobes.

The amino acid sequences for this enzyme have been determined for humans (20), horses (3), mice (22), *Trypanosoma brucei* (23), *S. cerevisiae* (26), and *Aspergillus nidulans* (6). These sequences share over 65% amino acid identity (6, 22). No information has been reported regarding the sequence of a bacterial phosphoglycerate kinase. Such information may provide further insight into essential structural features by defining regions which have been conserved in eucaryotes and procaryotes during evolution.

Phosphoglycerate kinase is an essential enzyme for ATP production in *Zymomonas mobilis*, an obligately fermentative bacterium (21). This organism produces only one net mole of ATP per mole of glucose consumed via the Entner-Doudoroff glycolytic pathway (21). Phosphoglycerate kinase activity is expressed at high levels in *Z. mobilis* and continues to increase in specific activity as cell growth slows and cultures enter stationary phase (24). This pattern of expression is similar for all but two of the glycolytic enzymes in *Z. mobilis*, pyruvate kinase and glyceraldehyde-3-phosphate dehydrogenase (24). It was somewhat surprising that phosphoglycerate kinase expression and glyceraldehyde-3-phosphate dehydrogenase expression were discordant, since the enzymes function together in glycolysis in a coupled reaction (14, 29) and have been proposed to exist as a complex in yeasts (1).

Previous studies from our laboratory have reported the cloning and sequencing of the gene encoding glyceraldehyde-3-phosphate dehydrogenase (*gap*) from *Z. mobilis* (10). During the current investigation, the promoterless gene encoding phosphoglycerate kinase (*pgk*) was found to lie immediately downstream from the *gap* gene. Thus, in *Z. mobilis*, the *pgk* gene is the second gene within the *gap* operon.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The strains and plasmids used in this study are shown in Table 1. *Z. mobilis* CP4 was grown at 30°C in complex medium (24) containing 20 g/liter of glucose. *Escherichia coli* DH5α and TC4 were grown in Luria broth (18) or on Luria agar (1.5% agar) lacking added carbohydrate. Strain DF576 (CGSC 5966) was grown in minimal medium as described by Thomson et al. (31). Carbohydrates were sterilized by filtration and added to autoclaved basal medium after cooling. Selections for the complementation of the phosphoglycerate kinase defect in strain DF576 were made on Luria agar lacking added carbohydrate and on minimal medium containing glucose (31) with appropriate antibiotics. Ampicillin (50 mg/liter) and chloramphenicol (40 mg/liter) were used to select recombinants of *E. coli*; 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (20 mg/liter) was used to identify DNA insertions which inactivated β-galactosidase activity in strain DH5α containing derivatives of pUC18.

Cloning of the *Z. mobilis* gene encoding phosphoglycerate kinase. A previously constructed library of *Z. mobilis* chromosomal DNA in pLOI193 (7) was used to complement the *pgk* defect in strain DF576. Clones containing the gene encoding phosphoglycerate kinase were selected for growth

* Corresponding author.
† Florida Agricultural Experiment Station publication no. 8363.

TABLE 1. Strains and plasmids used in this study

Bacterial strain or plasmid	Relevant genotype	Source or reference
Bacterial strains		
Z. mobilis CP4	Prototroph	24
E. coli		
DF576	*pgk*	31
DH5α	*lacZ*M15 *recA*	BRL[a]
TC4	*recA*	8
Plasmids		
pLOI193	*cat tet*	7
pUC18	*bla lacI'Z'*[b]	BRL
pLOI310	*gap pgk cat*	10
pLOI311	*gap bla*	10
pLOI323	*gap pgk cat*	This study
pLOI320	*pgk cat*	This study
pLOI324	*pgk bla*	This study
pLOI325	*pgk bla*	This study
pLOI322	*pgk bla*	This study

[a] BRL, Bethesda Research Laboratories.
[b] Incomplete *lacI* and incomplete *lacZ*.

on Luria agar plates and on glucose minimal medium containing chloramphenicol. Plasmids were isolated from positive clones and transferred to strain DH5α for further study. Transformation, restriction mapping, and subcloning were carried out by using standard methods (19).

Assay of phosphoglycerate kinase. Cells were harvested, disrupted, and assayed as previously described (24). Activities of phosphoglycerate kinase (EC 2.7.2.3) are expressed as international units per milligram of total cell protein. Protein was determined was described by Layne (17) with bovine serum albumin as a standard.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Gels were prepared and stained with Coomassie blue as previously described (9).

Determination of DNA sequence and mapping of transcriptional initiation. The *pgk* gene from *Z. mobilis* was sequenced in both directions by the dideoxy method as previously described (7). A 1.86-kilobase region was sequenced in both directions using DNA fragments generated by *Sau*3A and other restriction enzymes in derivatives of bacteriophages M13mp18 and M13mp19. This sequenced region began at the *Pst*I site between *gap* and *pgk* and continued for 467 bases below the *Hin*dIII site at the carboxy terminus of *pgk*. An oligonucleotide primer was used to sequence a portion of the amino terminus and upstream region (5'CTAATGTACGAAAAGCC3'). A second oligonucleotide primer was used to sequence a portion of the carboxy terminus (5'GCCGGGTGTTAAAGC3').

Sequence data were analyzed with the programs of Pustell and Kafatos (28), obtained from International Biotechnologies, Inc. (New Haven, Conn.). Protein alignment for homology was performed using the algorithm described by Wilbur and Lipman (32).

The 5' termini of transcripts were mapped by the primer extension method as previously described (8, 9) by using the oligonucleotide complementary to the noncoding strand spanning codons 2 through 6 at the N terminus.

Primers used for these experiments were synthesized with an Autogen 500 oligonucleotide synthesizer from Millipore/Genetic Design (Bedford, Mass).

Enzymes and chemicals. Restriction enzymes and DNA-modifying enzymes were obtained from Bethesda Research Laboratories, Inc. (Gaithersburg, Md.). M13 sequencing primer was obtained from New England BioLabs, Inc. (Beverly, Mass.). Radioactive compounds were purchased from New England Nuclear Corp. (Boston, Mass.). Biochemicals were purchased from Sigma Chemical Co. (St. Louis, Mo.).

RESULTS

Cloning of the gene encoding phosphoglycerate kinase. Plasmids containing our library of *Z. mobilis* DNA were transformed into strain DF576 with selection for growth on Luria agar containing chloramphenicol. Ten positive clones were isolated from an estimated 10,000 chloramphenicol-resistant colonies. Plasmid DNA was isolated from each of these original clones and retested to confirm its ability to restore growth ability of strain DF576 on glucose minimal medium. On the basis of digestions with *Bam*HI, *Pst*I, and *Eco*RV, these clones fell into three groups with insert sizes ranging from 4.7 kilobase pairs (kb) to over 6 kb. Two of these, designated PLOI320 (4.7 kb) and pLOI323 (5.7 kb), were chosen for further study (Fig. 1).

In PLOI320, both *Bam*HI sites were regenerated during the ligation of the *Sau*3A fragment from *Z. mobilis*. Both *Bam*HI sites were lost in the construction of pLOI323.

On the basis of restriction maps, pLOI320 and pLOI323 have a region of approximately 2.1 kb in common which was presumed to contain the *pgk* gene. The coding region for this gene was further localized by subcloning into pUC18, using strain DH5α (Fig. 1). The *Eco*RV fragment from pLOI323 was cloned into pUC18 in both orientations to construct pLOI324 and pLOI325 (results not shown). The *Bam*HI-to-*Hin*dIII plus *Hin*dIII-to-*Hin*dIII fragments (1.23 kb) from pLOI320 were cloned into appropriate polylinker sites of pUC18 to construct pLOI322. All of these constructions complemented the *pgk* defect in strain DF576 and allowed growth on glucose minimal medium and growth on Luria agar. Three other constructions were made by inserting fragments from pLOI320 into pUC18. These constructions did not complement the *pgk* defect in strain DF576 (Fig. 1). Our data indicate that the *pgk* gene is located within a 1.23-kb *Bam*HI-to-*Hin*dIII fragment containing internal *Pst*I, *Eco*RI, and *Hin*dIII sites.

The similarity of the restriction map for pLOI323 to that obtained for pLOI310 containing the *Z. mobilis gap* gene (10) suggested that *pgk* may be located on the same DNA fragment containing *gap*. This was confirmed by the ability of pLOI310 to complement strain DF576 and the inability of pLOI311, which contains the *gap* gene and only part of the region predicted to span the *pgk* gene, to do so.

Expression of *Z. mobilis* gene encoding phosphoglycerate kinase in *E. coli*. Table 2 summarizes the specific activities of phosphoglycerate kinase in strains of *E. coli* lacking and harboring plasmids. Strain DF576 exhibited very low levels of phosphoglycerate kinase activity as shown previously (31). Phosphoglycerate kinase activity was readily detected in strain TC4 which is wild type for this gene. The level of phosphoglycerate kinase in derivatives of strain DF576 which harbored plasmids shown to complement the *pgk* defect exceeded that found in strain TC4 (wild type for this gene). Good expression was obtained in both original clones, although pLOI320 lacked the native *gap* promoter. In this clone, expression is presumed to result from the *tet* promoter. The highest level of expression, 6.1 IU/mg of cell protein, was obtained in the original clone, pLOI323, which contained the native *Z. mobilis gap* promoter. A similar high

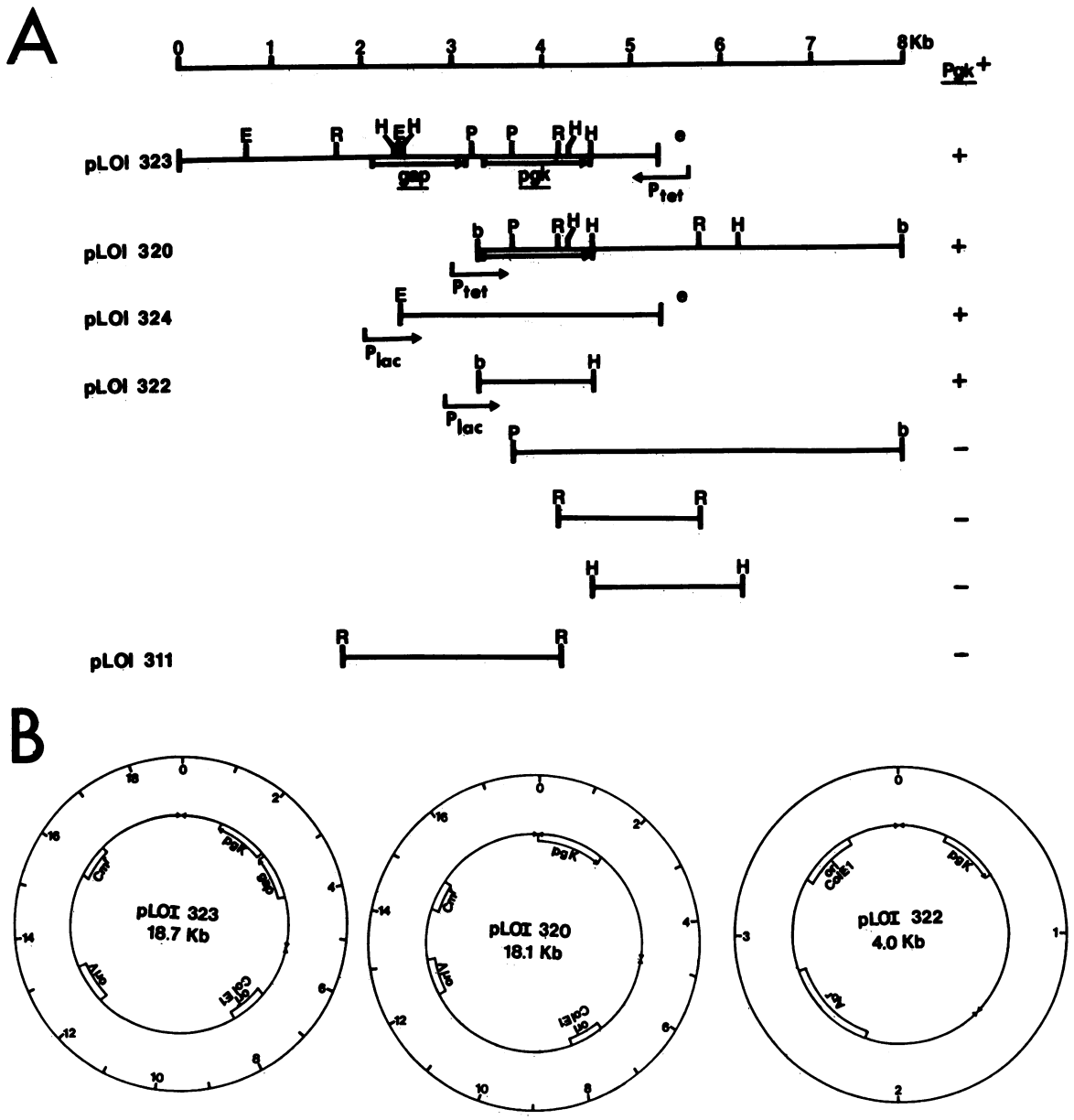

FIG. 1. Linear (A) and circular (B) restriction maps comparing the DNA fragments and gene locations in plasmids used in this study. Single uppercase letters refer to restriction enzyme sites in the *Z. mobilis* DNA, and single lowercase letters refer to restriction enzyme sites present in the vector. Arrows indicate direction of transcription. Plus or minus signs on the right side of restriction maps in panel A denote ability to complement the *pgk* mutation in *E. coli* DF576. Abbreviations: gap, *gap* gene; pgK, *pgk* gene; Ap^r, β-lactamase; Cm^r, chloramphenicol acyl transferase; b, *Bam*HI; H, *Hin*dIII; R, *Eco*RI; P, *Pst*I; E and e, *Eco*RV; P_{lac}, *lac* promoter; P_{tet}, *tet* promoter.

level of expression was observed in *E. coli* for glyceraldehyde-3-phosphate dehydrogenase with the *gap* promoter (10). Strains DF576(pLOI324) and DF576(pLOI325) contain the same *Z. mobilis* fragment with the *pgk* gene in opposite orientations with respect to the *lac* promoter and lack the native *Z. mobilis* promoter. Although the *Z. mobilis pgk* gene was expressed in both, threefold-higher activity was observed in DF576(pLOI324) with the *lac* promoter acting in the direction of *pgk* transcription compared with DF576 (pLOI325).

Pawluk et al. (25) have previously reported the molecular weight of *Z. mobilis* phosphoglycerate kinase as being approximately 44,000. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of solubilized cells revealed a band near

TABLE 2. Expression of *Z. mobilis* gene encoding phosphoglycerate kinase in *E. coli*

Strain	Plasmid	PGK activity[a]
DF576	None	0.01
TC4	None	0.62
DF576	pLOI323	6.1
DF576	pLOI320	2.9
DF576	pLOI324	4.9
DF576	pLOI325	1.3

[a] Specific activity (in international units per milligram of cell protein) of phosphoglycerate kinase.

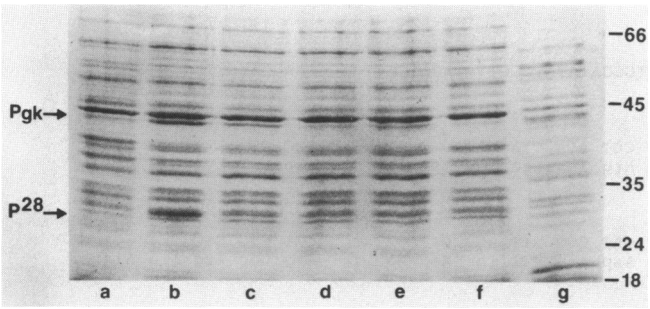

FIG. 2. Sodium dodecyl sulfate-polyacrylamide gel showing expression of the *Z. mobilis pgk* gene in *E. coli*. Lanes: a, strain DF576 without plasmid; b, strain DF576(pLOI323); c, strain DF576(pLOI324); d, DF576(pLOI325); e, DF576(pLOI320); f, strain TC4; g, *Z. mobilis* CP4. The position of the *Z. mobilis* phosphoglycerate kinase (Pgk) and an unidentified protein of approximately 28,000 in molecular weight (p28) are labeled on the left. The positions of molecular weight markers (in thousands) are indicated on the right.

this region (approximate molecular weight of 42,000) which was much more intense in DF576(pLOI323) than in DF576 lacking plasmid (Fig. 2). This band varied in intensity among derivatives of strain DF576 containing different plasmids with the *Z. mobilis pgk* gene. These differences in intensity were commensurate with the specific activities shown in Table 2, and this band is presumed to represent the *Z. mobilis* phosphoglycerate kinase. This band corresponds to one of the major bands in *Z. mobilis* CP4 and to a relatively faint band in *E. coli* TC4 (wild type for *pgk*).

The plasmid, pLOI323, also contains a functional *Z. mobilis gap* gene. Although the subunit molecular weight of this protein is 36,099 based upon sequence and gel filtration (25), this protein is reported to behave anomalously on sodium dodecyl sulfate gels with an apparent molecular weight of 42,000 (25). Thus on a sodium dodecyl sulfate-polyacrylamide gel, the products from both the *gap* and *pgk* genes appear in the same position.

An additional band of increased intensity with an apparent molecular weight of 28,000 was observed in DF576 (pLOI323). This protein is presumed to be encoded by unsequenced regions of the *Z. mobilis* DNA fragment above the *pgk* and *gap* genes.

Sequence of the phosphoglycerate kinase gene. The open reading frame which was observed downstream from the carboxy terminus of *Z. mobilis gap* (10) is presumed to correspond to the amino terminus of the *pgk* gene. The methionine at this position defines a 396-amino-acid open reading frame with considerable homology to known eucaryotic phosphoglycerate kinase sequences. The nucleotide and translated amino acid sequences for *pgk* from *Z. mobilis* are shown in Fig. 3. This sequence is shown as beginning at the *Pst*I site between *pgk* and *gap*, although the complete intervening region of 225 base pairs has been sequenced and reported previously (10). Two potential Shine-Dalgarno regions (GGAGG) were observed in this intervening region, followed by a start codon. However, the reading frame defined by the first of these terminated in the third codon. The second Shine-Dalgarno site is located five bases upstream from the open reading frame defining phosphoglycerate kinase. The region between translational initiation and the ribosomal-binding site contains only adenosine and thymine residues, consistent with efficient translation.

The *Z. mobilis pgk* gene contains an open reading frame encoding 396 amino acids (including the N-terminal methionine), with a molecular weight of 41,384. The reading frame is terminated by the codon, TAA. Six additional termination codons are present within the downstream 100 base pairs which cover all possible reading frames. This gene contains an unusually low level of tyrosine, consistent with the low A_{260} reported for the purified enzyme (25).

The intervening sequence between *pgk* and *gap* is particularly A+T rich and contains palindromic and repetitive sequences which could potentially form secondary structures affecting transcription and translation. There are at least three large palindromes potentially capable of forming a stem loop secondary structure. One of these is immediately downstream from the carboxy terminus of *gap*, TTTTTGC GAT/CTT/ATCGGTAAAA with two mismatches, as previously described (10). At 28 base pairs further downstream from *gap*, a second palindrome is present with a single mismatched position, CTTAATTACT/GCAGTT/AGTTAT TAAG (upstream from the sequence shown in Fig. 3). The third palindrome is immediately upstream from the translational initiation site of *pgk*, CCGCTTAT/GAT/ATAA GGCGG, with one mismatched position. A sequence (CCTCCT) complementary to the Shine-Dalgarno sequence was also observed beginning at base pair 35, between the second and third palindromes. These palindromic sequences could function to alter the efficiency of expression of *pgk* relative to the proximal gene, *gap*, or to protect this message from degradation by RNase.

The sequence distal to the carboxy terminus is unusual. No palindromes other than homopolymeric regions of T and A were observed. The sequence, CCTGCGA, occurred 100 base pairs distal to the *pgk* termination codon (TAA) and was repeated a total of 52 times. The significance of this repeated sequence (which appears much like a zipper on sequencing gels) is unknown.

Codon usage and codon bias. Table 3 shows a comparison of codon usage for *pgk* to the average codon usage for this gene and three other genes involved in ethanol production in *Z. mobilis*. The pattern of codon usage in *pgk* was very similar to the average codon usage of *Z. mobilis gap*, *pgk*, *adhB*, and *pdc*. Twelve codons were unused and nine were used only once in *pgk*. Nine codons were used no more than once by the combination of all four genes. Two codons, GTA (Val) and ACA (Thr), were used once in *pgk* and were unused in any of the other three genes. The four *Z. mobilis* genes were most highly biased in the their utilization of codons for arginine, cysteine, glycine, glutamic acid, glutamine, and phenylalanine.

The method of Bennetzen and Hall (4) was used to calculate codon bias for these four genes. All were similar, with a codon bias of 0.78 for *pgk*, 0.77 for *pdc*, 0.74 for *adhB*, and 0.73 for *gap* when calculated using amino acids exhibiting a high degree of codon preference. Only amino acids in which the sum of half the codons exceeded 75% of total usage for the combined four genes were considered highly biased and were used to calculate this parameter. This requirement eliminated Asn, Asp, His, Leu, Lys, Met, Ser, Trp, and Tyr from the calculation. The most frequently used codon set consists of TTC, CTG, ATC, ATG, GTT, TCC, CCG, ACC, GCT, TAT, CAC, CAG, AAC, AAA, GAT, GAA, TGC, TGG, CGT, and GGT.

Transcriptional initiation of *pgk* in *E. coli* and *Z. mobilis*. The initiation of transcription was investigated by primer extension analysis. The sequence, 5′CTAATGTACG AAAAGCC3′, which is complementary to the message between codons 2 through 6 in the N terminus was used as the primer. RNA from *E. coli* DH5α(pLOI323), from *Z.*

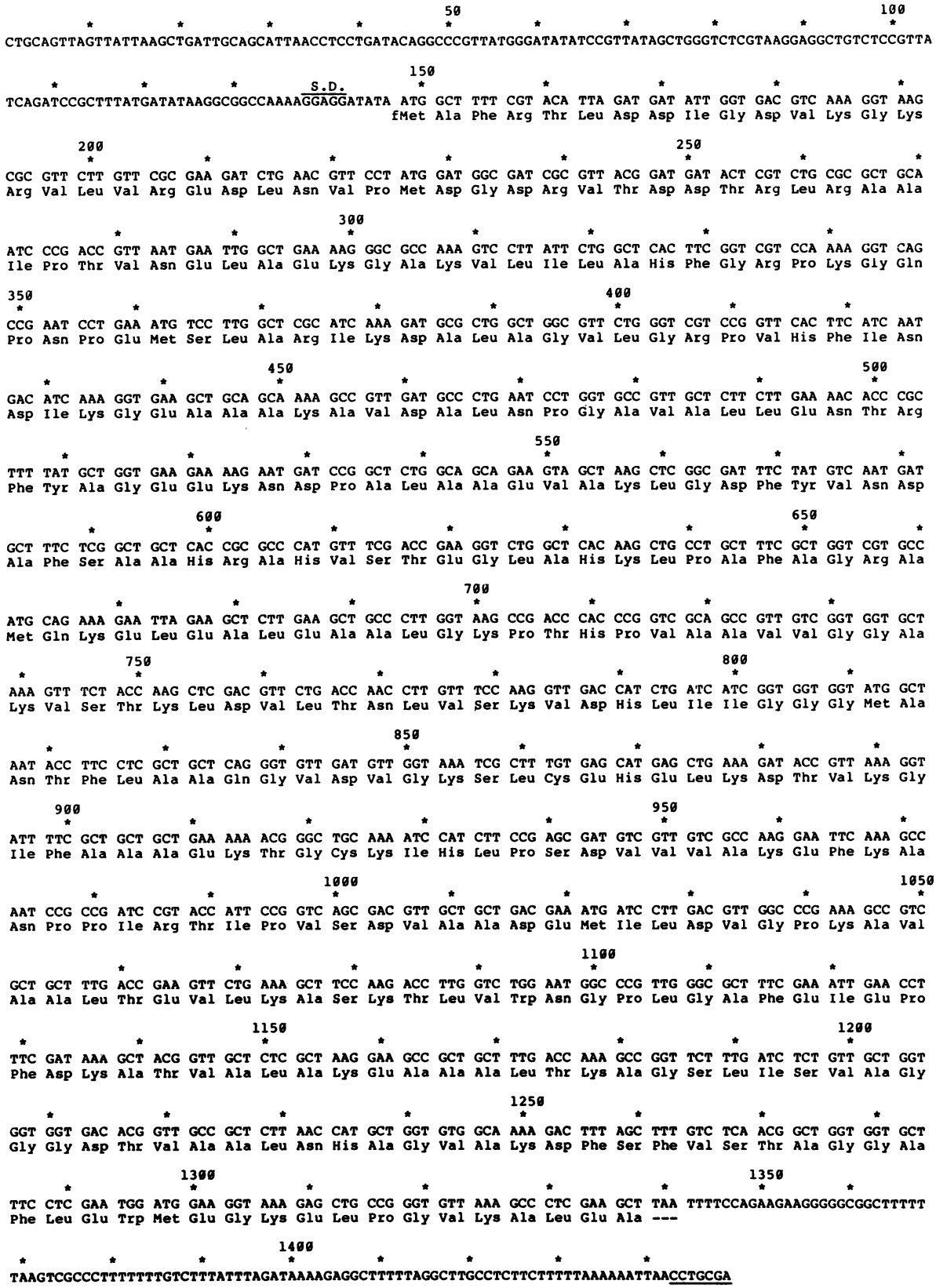

FIG. 3. Sequence and translation of the *Z. mobilis* gene encoding phosphoglycerate kinase. The proposed ribosomal binding site is overlined and labeled (S.D.). The seven underlined bases at the end of the sequence are repeated 52 times.

TABLE 3. Codon usage in *Z. mobilis*

Amino acid	Codon	Frequency		Amino acid	Codon	Frequency	
		PGK[a]	Combined[b]			PGK	Combined
Phe	TTT	4	9	Ala	GCT	43	138
	TTC	11	44		GCC	15	51
					GCA	7	40
Leu	TTA	2	4		GCG	1	14
	TTG	7	25				
	CTT	11	32	Tyr	TAT	2	26
	CTC	6	28		TAC	0	12
	CTA	0	0				
	CTG	14	65	His	CAT	5	18
					CAC	5	21
Ile	ATT	5	21				
	ATC	10	57	Gln	GAA	0	2
	ATA	0	0		GAG	3	27
Met	ATG	7	45	Asn	AAT	9	26
					AAC	4	52
Val	GTT	26	94				
	GTC	11	43	Lys	AAA	21	66
	GTA	1	1		AAG	11	45
	GTG	1	9				
				Asp	GAT	16	57
Ser	TCT	3	19		GAC	9	44
	TCC	3	23				
	TCA	1	5	Glu	GAA	23	86
	TCG	3	6		GAG	3	6
	AGT	0	4				
	AGC	3	21	Cys	TGT	1	3
					TGC	1	15
Pro	CCT	4	14				
	CCC	0	4	Trp	TGG	2	10
	CCA	1	6				
	CCG	13	49	Arg	CGT	6	33
					CGC	7	22
Thr	ACT	1	10		CGA	0	0
	ACC	12	64		CGG	0	1
	ACA	1	1		AGA	0	0
	ACG	5	21		ACG	0	0
				Gly	GGT	28	97
					GGC	8	38
					GGA	0	2
					GGG	0	0

[a] *Z. mobilis* phosphoglycerate kinase.
[b] Total for *Z. mobilis gap*, *pgk*, *adhB*, and *pdc*.

mobilis CP4, and from *E. coli* DH5α (lacking plasmid, negative control) were used as templates. In both *Z. mobilis* CP4 and *E. coli* DH5α(pLOI323), a prominent band (data not shown) was observed very high on the sequencing gel at a position over 1,000 base pairs upstream from the translational start of *pgk*. This band was absent in *E. coli* DH5α lacking plasmid. These results indicate that no transcriptional start exists between the *Z. mobilis pgk* and *gap* genes. In addition, no regions were located which exhibited strong homology to the transcriptional regions previously identified in *Z. mobilis gap*, *adhB*, *pdc* or the proposed consensus sequence for *Z. mobilis* (10). Thus, the *Z. mobilis pgk* gene is expressed as part of a transcriptional unit which includes the *gap* gene, designated the *gap* operon.

Conservation of primary amino acid sequence among genes encoding phosphoglycerate kinase. *Z. mobilis* is the first bacterium for which the primary sequence of phosphoglycerate kinase has been determined. This sequence was compared with the known eucaryotic sequences (Table 4) and revealed a high degree of amino acid conservation between these evolutionarily diverse organisms. From 173 to 185 amino acid positions were shared between *Z. mobilis* and each of the other organisms. Although the extent of amino acid identity was very similar in all comparisons, the *Z. mobilis* sequence is most similar to that recently reported for *A. nidulans* (6).

The three-dimensional structure of phosphoglycerate kinase is known for *S. cerevisiae* (26), and the *Z. mobilis* primary amino acid sequences corresponding to these structural features can be inferred by comparison (Fig. 4). Since the eucaryotic sequences are even more highly conserved than that of *Z. mobilis*, a comparison of the amino acid positions which are maintained in all known sequences should identify regions which are most critical as defined by their having been conserved during evolution. These positions are indicated in Fig. 4 by the solid circles, with 102 amino acid positions being conserved in all sequences. The most highly conserved region was at the carboxy terminus including the final two α-helical regions (positions 378 through 386 and 390 through 395) and connecting amino acids. This region forms part of the hinge between the two lobes of phosphoglycerate kinase and is also part of the binding site for bis-phosphoglycerate. In general, the amino acids at the ends of secondary structures and the connecting amino acids between organized secondary structures are most highly conserved, followed by β-sheets. As observed previously (22), the turns and portions of the β-sheets which form the substrate-binding sites are highly conserved.

DISCUSSION

The gene encoding phosphoglycerate kinase appears to be part of an operon containing at least two glycolytic genes in *Z. mobilis*. Our sequencing data and mapping of transcriptional initiation indicate that the promoter immediately upstream from the *gap* gene also serves *pgk*. Although the products of these two genes function together to generate ATP during glycolysis and may form a complex in vivo (1), this is the first report that these two genes are genetically linked in an operon. Some of the genes in glycolysis are clustered in *E. coli* (2, 14) and in *Pseudomonas* spp. (12, 16), such as the combination of *edd*, *eda*, and *zwf*. However, *gap* and *pgk* are quite distant in the genetic map of *E. coli* (2), and their locations have not been reported in *Pseudomonas* spp. (16).

Z. mobilis is obligately fermentative and utilizes very few sugars, all of which are metabolized via the Entner-Doudoroff pathway (21). Our previous studies have shown that during growth and fermentation, the intracellular activities of phosphoglycerate kinase and glyceraldehyde-3-phosphate dehydrogenase do not change in a coordinate manner. The specific activity of the upstream enzyme nearest the pro-

TABLE 4. Amino acid identity between *Z. mobilis* phosphoglycerate kinase and that of eucaryotic organisms

Organism	% Identity	Reference
Aspergillus nidulans	46.6	7
Saccharomyces cerevisiae	44.8	26
Human	44.8	20
Trypanosoma brucei, allele C	44.3	23
Mouse	43.8	22
Horse	43.6	3
Trypanosoma brucei, allele B	43.6	23

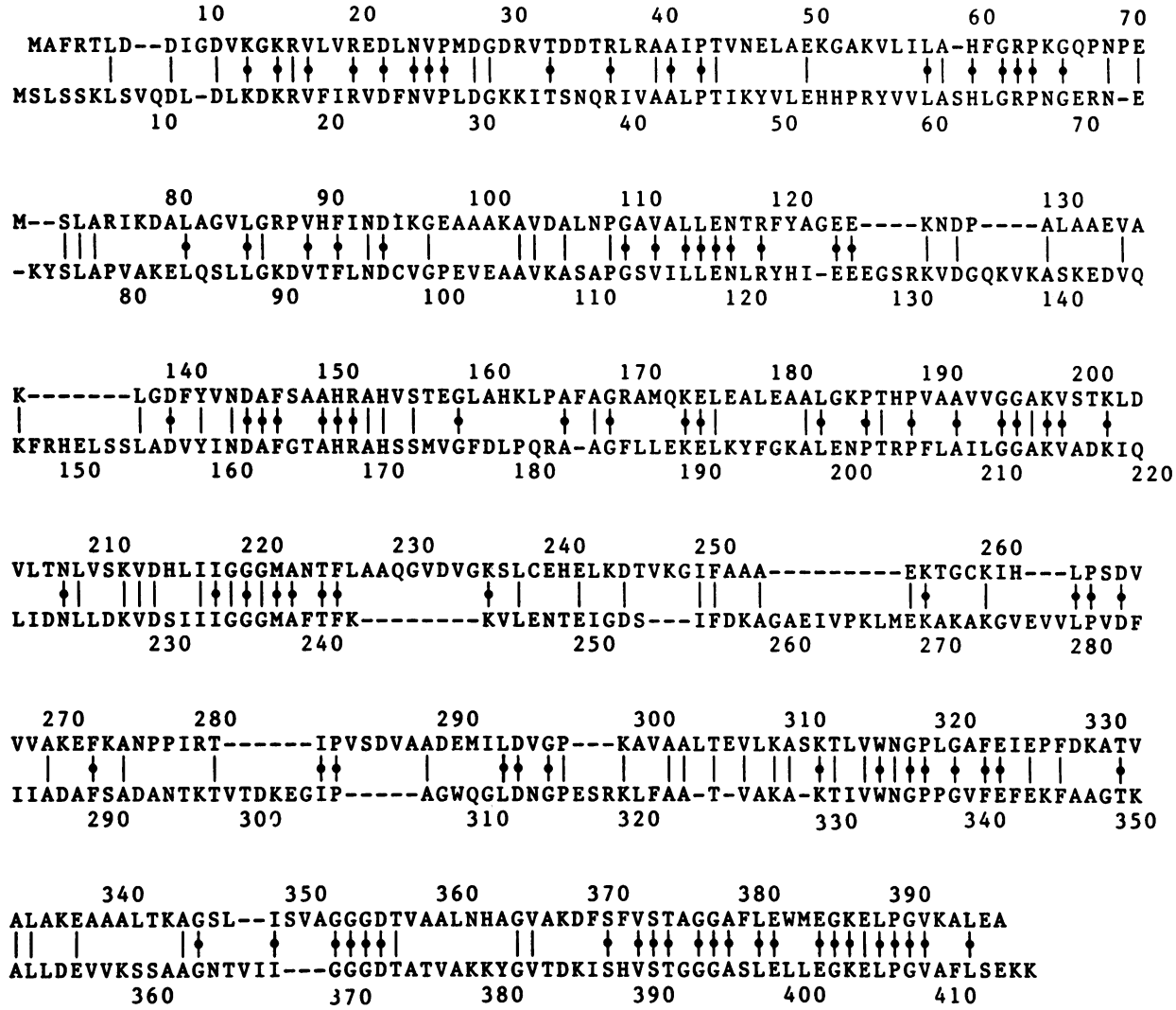

FIG. 4. Comparison of amino acid homology between *gap* genes. Identity between the *Z. mobilis* gene and that of *S. cerevisiae* (|) and identity shared by all seven sequences of phosphoglycerate kinase (ϕ) are shown.

moter, glyceraldehyde-3-phosphate dehydrogenase, declines in specific activity, whereas that of phosphoglycerate kinase continues to rise during the transition from exponential to stationary phase (24). Together, these results provide evidence that differential control of these two genes is not at the level of transcription. The apparent uncoupling of expression may be due to differences in the pH stability of these enzymes, since internal pH begins to fall under these conditions (24). Alternatively, this difference in expression could reflect differences in translational efficiency or mRNA stability. The intervening sequence connecting these two genes contains two potential ribosomal binding sites and at least three palindromic sequences capable of a stem and loop conformation. In addition, the sequence downstream from phosphoglycerate kinase contains an unusual 52-fold repeat of a 7-base sequence, the zipper sequence. The double ribosomal binding sites may enhance translation. The positioning of the coding region for *pgk* between the unusual downstream zipper sequence and a highly structured intervening sequence with *gap* attached may serve as protection from RNase activity and allow further translation during the transition from active growth to stationary phase.

The *pgk* gene and the three other *Z. mobilis* genes (9–11) involved in fermentation all exhibit strong codon bias, as would be expected for highly expressed genes (13). All of these genes contain A+T-rich regions immediately surrounding Shine-Dalgarno sites conducive to efficient ribosomal attachment and translation (30). It is likely that the combined codon usage by these four highly expressed genes reflects the relative abundance of the different tRNA species (13).

A comparison of the amino acid sequences for phosphoglycerate kinase from *Z. mobilis* and a variety of eucaryotes indicates that a high degree of constraint has been imposed by function during evolution. Considerable amino acid identity and a much higher level of functional homology (if conservative replacements by chemically similar amino acids are considered) has been retained between eucaryotic and bacterial enzymes. Based upon the similarities in amino acid sequence with yeasts, it is likely that the *Z. mobilis* enzyme assumes a similar conformation. The most highly conserved regions were the borders of secondary structural features and the connecting amino acids between these organized regions. Such conservation would tend to pre-

serve the folding patterns of the enzyme while permitting a variety of substitutions which cause only minimal changes in α-helical regions or in β-sheets. Comparison among all sequences indicated that the substrate-binding regions were highly conserved as noted previously (22) and that the terminal segments of the carboxy terminus which formed part of the hinge were also highly conserved.

ACKNOWLEDGMENTS

We appreciate the help of B. Bachmann in providing mutant strains of *E. coli* and advice concerning their handling. We also thank K. T. Shanmugam and R. A. Jensen for their suggestions during the preparation of this manuscript.

These studies were supported in part by grants from the National Science Foundation (DMB 8204928), by the Department of Energy, Office of Basic Energy Sciences (DE-FG05-86ER13575), by the U.S. Department of Agriculture, Alcohol Fuels Program (86-CRCR-1-2134), and by the Florida Agricultural Experiment Station.

LITERATURE CITED

1. **Ashmarina, L. I., V. I. Muronetz, and N. K. Nagradova.** 1985. Yeast glyceraldehyde-3-phosphate dehydrogenase. Evidence that subunit cooperativity in catalysis can be controlled by the formation of a complex with phosphoglycerate kinase. Eur. J. Biochem. **149**:67–72.
2. **Bachmann, B. J.** 1983. Linkage map of *Escherichia coli* K-12, edition 7. Microbiol. Rev. **47**:180–230.
3. **Banks, R. D., C. C. F. Blake, P. R. Evans, R. Haser, D. W. Rice, G. W. Hardy, M. Merrett, and A. W. Phillips.** 1979. Sequence, structure and activity of phosphoglycerate kinase: a possible hinge-bending enzyme. Nature (London) **279**:773–777.
4. **Bennetzen, J. L., and B. D. Hall.** 1982. Codon selection in yeast. J. Biol. Chem. **257**:3026–3031.
5. **Bresolin, N., A. Miranda, H. W. Chang, S. Shanske, and S. Dimauro.** 1984. Phosphoglycerate kinase deficiency myopathy: biochemical and immunological studies with mutant enzyme. Muscle & Nerve **7**:542–551.
6. **Clements, J. M., and C. F. Roberts.** 1986. Transcription and processing signals in the 3-phosphoglycerate kinase (PGK) gene from *Aspergillus nidulans*. Gene **44**:97–105.
7. **Conway, T., M. O.-K. Byun, and L. O. Ingram.** 1987. Expression vector for *Zymomonas mobilis*. Appl. Environ. Microbiol. **53**:235–241.
8. **Conway, T., Y. A. Osman, and L. O. Ingram.** 1987. Gene expression in *Zymomonas mobilis*: promoter structure and identification of membrane anchor sequences forming functional LacZ' fusion proteins. J. Bacteriol. **169**:2327–2335.
9. **Conway, T., Y. A. Osman, J. I. Konnan, E. M. Hoffmann, and L. O. Ingram.** 1987. Promoter and nucleotide sequences of the *Zymomonas mobilis* pyruvate decarboxylase. J. Bacteriol. **169**:949–954.
10. **Conway, T., G. W. Sewell, and L. O. Ingram.** 1987. Glyceraldehyde-3-phosphate dehydrogenase from *Zymomonas mobilis*: cloning, sequencing, and identification of promoter region. J. Bacteriol. **169**:5653–5662.
11. **Conway, T., G. W. Sewell, Y. A. Osman, and L. O. Ingram.** 1987. Cloning and sequencing of the alcohol dehydrogenase II gene from *Zymomonas mobilis*. J. Bacteriol. **169**:2591–2597.
12. **Cuskey, S. M., and P. V. Phibbs, Jr.** 1985. Chromosomal mapping of mutations affecting glycerol and glucose catabolism in *Pseudomonas aeruginosa* PAO. J. Bacteriol. **162**:872–880.
13. **de Boer, H. A., and R. A. Kastelein.** 1986. Biased codon usage: an exploration of its role in optimization of translation. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, Stoneham, Mass.
14. **Fraenkel, D. G., and R. T. Vinopal.** 1973. Carbohydrate metabolism in bacteria. Annu. Rev. Microbiol. **27**:69–100.
15. **Fujii, H., W. K. G. Keietsch, and A. Yoshida.** 1980. A single amino acid substitution (Asp→Asn) in a phosphoglycerate kinase variant (PGK Munchen) associated with enzyme deficiency. J. Biol. Chem. **255**:6421–6423.
16. **Holloway, B. W., and A. F. Morgan.** 1986. Genome organization in *Pseudomonas*. Annu. Rev. Microbiol. **40**:79–105.
17. **Layne, E.** 1957. Spectrophotometric and turbidometric methods for measuring proteins. Methods Enzymol. **3**:447–454.
18. **Luria, S. E., and M. Delbruck.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28**:491–511.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Michelson, A. M., C. C. F. Blake, S. T. Evans, and S. H. Orkin.** 1985. Structure of the human phosphoglycerate kinase gene and the intron-mediated evolution and dispersal of the nucleotide-binding domain. Proc. Natl. Acad. Sci. USA **82**:6965–6969.
21. **Montenecourt, B. S.** 1985. *Zymomonas*, a unique genus of bacteria, p. 261–289. *In* A. L. Demain and N. A. Solomon (ed.), Biology of industrial microorganisms. Benjamin-Cummings Publishing Co., Menlo Park, Calif.
22. **Mori, N., J. Singer-Sam, and A. D. Riggs.** 1986. Evolutionary conservation of the substrate-binding cleft of phosphoglycerate kinases. FEBS Lett. **204**:313–317.
23. **Osinga, K. A., B. W. Swinkels, W. C. Gibson, P. Borst, G. H. Veeneman, J. H. Van Boom, P. A. Michels, and F. R. Opperdoes.** 1985. Topogenesis of microbody enzymes. A sequence comparison of the genes for the glycosomal (microbody) and cytosolic phosphoglycerate kinases of *Trypanosoma brucei*. EMBO J. **4**:3811–3817.
24. **Osman, Y. A., T. Conway, S. J. Bonetti, and L. O. Ingram.** 1987. Glycolytic flux in *Zymomonas mobilis*: enzyme and metabolite levels during batch fermentation. J. Bacteriol. **169**:3726–3736.
25. **Pawluk, A., R. K. Scopes, and K. Griffiths-Smith.** 1986. Isolation and properties of the glycolytic enzymes from *Zymomonas mobilis*. Biochem. J. **238**:275–281.
26. **Perkins, R. E., S. C. Conroy, B. Dunbar, L. A. Fothergill, M. F. Tuite, M. J. Dobson, S. M. Kingsman, and A. J. Kingsman.** 1983. The complete amino acid sequence of yeast phosphoglycerate kinase. Biochem. J. **211**:199–218.
27. **Piper, P. W., B. Curran, M. W. Davies, A. Lockheart, and G. Reid.** 1986. Transcription of the phosphoglycerate kinase gene of *Saccharomyces cerevisiae* increases when fermentative cultures are stressed by heat-shock. Eur. J. Biochem. **161**:525–531.
28. **Pustell, J., and F. C. Kafatos.** 1984. A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis, and homology determination. Nucleic Acids Res. **12**:643–655.
29. **Scopes, R. K.** 1973. 3-Phosphoglycerate kinase, p. 335–351. *In* P. Boyer (ed.), The enzymes, vol. 8, 3rd ed. Academic Press, Inc., New York.
30. **Stormo, G. D.** 1986. Translation initiation, p. 195–224. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, Stoneham, Mass.
31. **Thomson, J., P. D. Gerstenberger, D. E. Goldberg, E. Gociar, A. O. de Silva, and D. G. Fraenkel.** 1979. ColE1 hybrid plasmids for *Escherichia coli* genes of glycolysis and the hexose monophosphate shunt. J. Bacteriol. **137**:502–506.
32. **Wilbur, W. J., and D. J. Lipman.** 1983. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA **80**:726–730.
33. **Wilson, C. A. B., N. Hardman, L. A. Fothergill-Gilmore, S. J. Gamblin, and H. C. Watson.** 1987. Yeast phosphoglycerate kinase: investigation of catalytic function by site-directed mutagenesis. Biochem. J. **241**:609–614.