JOURNAL OF BACTERIOLOGY, July 1993, p. 3926–3933
0021-9193/93/133926-08$02.00/0

Vol. 175, No. 13

Cloning, Sequencing, and Expression of the *Zymomonas mobilis* Phosphoglycerate Mutase Gene (*pgm*) in *Escherichia coli*†

LORRAINE P. YOMANO,[1] ROBERT K. SCOPES,[2] AND L. O. INGRAM[1]*

Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida 32611,[1] and Centre for Protein and Enzyme Technology, La Trobe University, Bundoora, Victoria, Australia 3083[2]

Received 26 January 1993/Accepted 22 April 1993

Phosphoglycerate mutase is an essential glycolytic enzyme for *Zymomonas mobilis*, catalyzing the reversible interconversion of 3-phosphoglycerate and 2-phosphoglycerate. The *pgm* gene encoding this enzyme was cloned on a 5.2-kbp DNA fragment and expressed in *Escherichia coli*. Recombinants were identified by using antibodies directed against purified *Z. mobilis* phosphoglycerate mutase. The *pgm* gene contains a canonical ribosome-binding site, a biased pattern of codon usage, a long upstream untranslated region, and four promoters which share sequence homology. Interestingly, *adhA* and a D-specific 2-hydroxyacid dehydrogenase were found on the same DNA fragment and appear to form a cluster of genes which function in central metabolism. The translated sequence for *Z. mobilis pgm* was in full agreement with the 40 N-terminal amino acid residues determined by protein sequencing. The primary structure of the translated sequence is highly conserved (52 to 60% identity with other phosphoglycerate mutases) and also shares extensive homology with bisphosphoglycerate mutases (51 to 59% identity). Since Southern blots indicated the presence of only a single copy of *pgm* in the *Z. mobilis* chromosome, it is likely that the cloned *pgm* gene functions to provide both activities. *Z. mobilis* phosphoglycerate mutase is unusual in that it lacks the flexible tail and lysines at the carboxy terminus which are present in the enzyme isolated from all other organisms examined.

Phosphoglycerate mutase (PGM) functions in the lower segment of both the Entner-Doudoroff and the Embden-Meyerhof-Parnas glycolytic pathways. This enzyme is essential for *Zymomonas mobilis*, an obligately fermentative gram-negative bacterium, as it catalyzes the reversible conversion of 3-phosphoglycerate (3-PGA) to 2-PGA. Although the genes encoding 2,3-bisphosphoglycerate (BPGA)-dependent PGM have been extensively investigated in eukaryotes (16), only a single prokaryotic gene has been reported, that of *Streptomyces coelicolor* (38). The primary sequences of eukaryotic and *S. coelicolor* PGM are highly conserved, encoding proteins with M_rs of 26,000 to 27,000. Considerable information is available concerning the three-dimensional structure and the kinetic mechanism of this enzyme (36, 39). PGM also functions in many organisms as a 2,3-BPGA mutase, a 2,3-BPGA synthase, and a PGA phosphatase (36, 37). In some organisms, a separate BPGA mutase which shares extensive homology with PGM and may be considered an isoenzyme (13). is also present. Phosphorylation of the PGM active site by BPGA is essential for mutase activity.

In *Z. mobilis*, PGM is not under allosteric control but does exhibit a requirement for a low level of BPGA to maintain phosphorylation (27). The active form is a dimer with a subunit M_r of 26,000 and a specific activity of 2,000 IU/mg of protein. This enzyme is resolved into two spots on two-dimensional protein gels which may correspond to the phosphorylated and unphosphorylated forms. Together, these represent approximately 1% of the total cellular protein (2).

In this study, we have cloned and sequenced the *Z. mobilis* gene encoding PGM (*pgm*) on a 5.2-kbp DNA fragment. This fragment contains three additional open reading frames (ORFs) and the N terminus of the *adhA* gene (21).

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. Strains and plasmids used in this study are summarized in Table 1. *Z. mobilis* CP4 was grown at 30°C in complex medium (26) containing 10% glucose. Stock cultures were maintained on solid medium containing 2% glucose and 1.5% agar. *Escherichia coli* strains were grown at 37°C in Luria broth (31) or on Luria agar (1.5% agar). Recombinant *E. coli* DH5α and TC4 containing cloned *pgm* genes were supplemented with 2% glycerol and 50 mg of ampicillin per liter. Growth was monitored with a Spectronic 70 spectrophotometer (Bausch & Lomb, Inc., Rochester, N.Y.) at 550 nm.

DNA manipulations. Standard methods were used for the purification of DNA, plasmid construction, and the transformation of *E. coli* (31). Digestions with restriction enzymes were carried out as recommended by the manufacturers.

Construction of a *Z. mobilis* genomic DNA library. Chromosomal DNA was isolated from *Z. mobilis* as described previously (5) and used to construct a *Sau*3A library of 4- to 6-kbp fragments in pUC18. Clones were selected for ampicillin resistance with *E. coli* TC4 as the host. Approximately 2,000 colonies were pooled from the original transformation plates and used to isolate plasmid DNA. This DNA serves as an amplified genomic library.

Cloning of the PGM gene. *E. coli* TC4 was transformed with the genomic library. Nitrocellulose filter blots of transformants were screened for *pgm* expression by standard

* Corresponding author.

† Publication R-02895 from the University of Florida Agricultural Experiment Station.

TABLE 1. Strains and plasmids used

Strain or plasmid	Genetic characteristic(s)	Source or reference
Strains		
Z. mobilis CP4	Prototroph	26
E. coli		
TC4	*recA*	9
DH5α	Δ*lacZ*M15 *recA*	BRL[a]
S17-1 (λ*pir*)	*thi pro hsdR::RP4 recA*	14
Plasmids		
pUC18	*bla lacI'Z'*[b]	BRL
pBluescript II KS+	*bla lacZ*[c]	Stratagene
pLOI706EH	*tet lacI*^{q} *tacP*	3
pLOI707EH	*bla tet lacI*^{q} *tacP*	3
pLOI680[d]	*bla lacI'Z'*[b] *pgm ddh adhA'*	This study
pLOI706EH/*pgm*	pLOI706 containing the *pgm* ORF	This study

[a] BRL, Bethesda Research Laboratories, Gaithersburg, Md.
[b] Incomplete *lacI* and incomplete *lacZ*.
[c] Incomplete *lacZ*.
[d] Derivatives of pLOI680 are illustrated in Fig. 1.

methods (31) with rabbit anti-PGM antibody. Goat anti-rabbit immunoglobulin G conjugated to alkaline phosphatase (Sigma Chemical Co., St. Louis, Mo.) was used as the secondary antibody and developed with 5-bromo-4-chloro-3-indolylphosphate and Nitro Blue Tetrazolium. Positive clones were verified by enzyme assay.

Immunological methods. PGM was purified to apparent homogeneity as described previously (27). Polyclonal antibody to this preparation was prepared in New Zealand White rabbits and purified as previously described (1). Ouchterlony diffusion tests confirmed the presence of a single reactive band with *Z. mobilis* extracts and the absence of reactive bands with *E. coli* extracts. This antibody effectively neutralized *Z. mobilis* PGM activity and stained a single dominant band in Western blots (immunoblots) of *Z. mobilis* proteins (1). By using a 1:10,000 dilution of purified antibody, 40 pg of *Z. mobilis* PGM could be detected.

N-terminal amino acid sequence of *Z. mobilis* PGM. The N-terminal amino acid sequence of purified *Z. mobilis* PGM (27) was determined with an Applied Biosystems model 470A gas-liquid phase sequencer. Sequencing was unambiguous for the first 40 residues.

Protein extracts. Soluble extracts of *Z. mobilis* and *E. coli* DH5α were prepared from exponential-phase cells (optical density at 550 nm, 0.5) as described previously (2). These protein extracts were used for PGM enzyme assays, Ouchterlony analysis, and sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

PGM enzyme activity. PGM activity was measured as described by Pawluk et al. (27), with the addition of 0.1 mM 2,3-BPGA. Specific activity is expressed as international units per milligram of cellular protein (whole cells).

DNA sequencing and analysis. Double-stranded plasmid DNA was purified for sequencing by using Magic Mini Prep columns (Promega Corp., Madison, Wis.). The entire 5.2-kbp DNA insert in pLOI680 was sequenced by the dideoxy method of Sanger et al. (32) after subcloning of the *Pst*I-to-*Sma*I (polylinker sites) fragment into the corresponding sites of pBluescript II KS+ (Stratagene, La Jolla, Calif.). A series of ordered deletions were constructed by using the Erase-a-Base system (Promega Biotec, Madison, Wis.). The sequence was determined in both directions with Sequenase version 2.0 (United States Biochemical Corp., Cleveland, Ohio). Pyrophosphatase was added to the diluted Sequenase, and 7-deaza-dGTP extension mix was added to the termination mix (also 7-deaza-dGTP) as recommended by the manufacturer. Sequencing was carried out on a LI-COR (Lincoln, Nebr.) model 4000 DNA sequencer with fluorescent primers (25) and 7% acrylamide gels (Long Ranger gel concentrate; ATBiochem, Malvern, Pa.). Resulting sequences were analyzed with the Genetics Computer Group sequence analysis software package (University of Wisconsin, Madison), the National Center for Biotechnology Information data base searches using the BLAST network server (National Library of Medicine via INTERNET), and the Whitehead Institute sequence analysis programs (Massachusetts Institute of Technology, Cambridge).

Southern hybridization. A 642-bp *Hpa*I fragment within the *pgm* coding region (*Hpa*I fragment) was used to synthesize a digoxigenin-labeled probe (Genius DNA labeling and detection kit; Boehringer Mannheim Biochemicals, Indianapolis, Ind.). This probe was used to identify *pgm*-containing fragments in genomic digestions of *E. coli* DH5α and *Z. mobilis* CP4. Approximately 10 μg of digested DNA was loaded in each lane.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Gels (12% acrylamide) were prepared and stained with Coomassie blue, as described previously (21). Approximately 20 μg of cellular protein or 2 μg of purified PGM was loaded into each lane. Protein standards were obtained from Sigma Chemical Co.

Primer extension. The 5′ end of the *pgm* transcript was mapped by primer extension analysis (29). Total *Z. mobilis* CP4 RNA was isolated from mid-log-phase cells grown in 10% glucose as described previously (11). A primer (5′-CGGACTGTCCGTGACGGGAC-3′) complementary to the noncoding strand of the *pgm* gene was synthesized, labeled with ^{32}P, and hybridized to 100 μg of *Z. mobilis* RNA. The hybridization mixture was denatured for 5 min at 65°C and then incubated for 1 h at 42°C. A sequencing ladder was prepared with the same primer and pLOI680 (Fig. 1) as the DNA template.

Construction of *pgm* shuttle vector. The minimal coding region for *Z. mobilis* PGM was synthesized by using the polymerase chain reaction (30) with plasmid pLOI680 (Fig. 1) as a template. Reactions were performed with the TempCycler model 50 (Coy Laboratory Products, Inc., Ann Arbor, Mich.) and a GeneAmp kit (Perkin-Elmer Cetus, Norwalk, Conn.) containing *Taq* polymerase. Amplification reactions (30 cycles) contained 10 mM each deoxynucleoside triphosphate, 50 pmol of each primer, 2 ng of plasmid template, 1 μl of Perfect Match DNA polymerase enhancer (Stratagene), 1.5 mM $MgCl_2$, and 2.5 U of *Taq* polymerase in a total volume of 100 μl. The primer complementary to the 5′ end of *pgm* (5′-GCGCGAGCTCAGGAGTGGAATATGC CCA-3′) included the native ribosome-binding site (Shine-Dalgarno sequence, underlined) and a unique 5′ *Sst*I site. The primer complementary to the 3′ end of *pgm* (5′-GTGT GCGGCCGCGGCTATTAACGTTCG-3′) included a unique 3′ *Not*I site. These sites allowed directional cloning into the *lacI*^{q}-controlled expression vector pLOI707EH. The resulting plasmid, pLOI697, was conjugated into *Z. mobilis* as previously described (3).

Carbon flux studies. Overnight cultures of *Z. mobilis* CP4 recombinants were diluted into stoppered bottles containing 800 ml of complex medium (10% glucose) to provide an

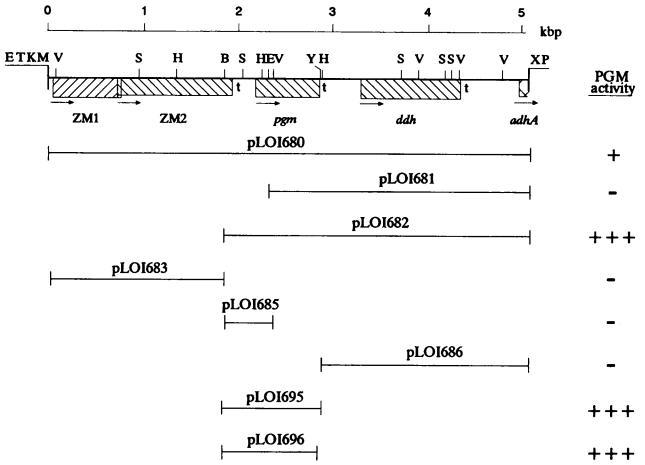

FIG. 1. Restriction map of *Z. mobilis* DNA fragment in pLOI680 and subclones. All clones are derivatives of pUC18. Hatched regions indicate ORFs. ZM1 and ZM2 overlap. Probable terminators (t) were identified by sequence analysis. Polylinker sites used in constructions are included on the ends. Subclones are aligned beneath restriction sites used for construction. Relative PGM activities are indicated at the right. Abbreviations: B, *Bgl*II; E, *Eco*RI; H, *Hpa*I; K, *Kpn*I; P, *Pst*I; M, *Sma*I; S, *Stu*I; T, *Sst*I; V, *Eco*RV; X, *Xba*I; Y, *Sty*I (additional *Sty*I sites are also present on pLOI680 outside the *pgm* coding region but are not shown). Arrows indicate direction of transcription.

initial optical density at 550 nm of 0.1 to 0.2. Cultures were grown with and without 4 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) to provide partial control of *pgm* expression. Inocula for induced cultures were also grown with IPTG. Glucose flux was estimated by measuring the rate of CO_2 evolution as previously described (3). The rate of CO_2 evolution is expressed in micromoles of CO_2 per minute per milligram of protein (international units per milligram of protein).

Nucleotide sequence accession numbers. The sequences for *pgm*, *ddh*, and the two unidentified ORFs (ZM1 and ZM2) have been assigned GenBank accession numbers LO9650, LO9651, and LO9649, respectively.

RESULTS

Cloning of the gene encoding PGM (*pgm*). Thirty clones expressing *Z. mobilis* CP4 PGM were isolated from plates containing a total of approximately 6,650 recombinants. On the basis of restriction mapping with a variety of endonucleases, these 4- to 6-kbp inserts fell into nine groups. The clone expressing the highest level of activity, DH5α (pLOI680), was selected for further study (Fig. 1). Ouchterlony analysis showed identity between single precipitin bands in protein extracts from this clone, in extracts from *Z. mobilis*, and in purified *Z. mobilis* PGM (data not shown).

A variety of subclones were used to identify the coding region for *pgm* (Fig. 1). Subclone pLOI681 was constructed by deleting the *Eco*RI (one site in the polylinker) DNA fragment and did not retain PGM activity. Subclone pLOI683 was constructed by deleting the *Bgl*II-to-*Xba*I (polylinker) DNA fragment and also lacked PGM activity. Subclone pLOI682 was constructed by deleting the *Bgl*II-to-*Kpn*I (polylinker site) DNA fragment and expressed a higher level of activity than the parental plasmid, pLOI680. Subclone pLOI685 was constructed by deleting the *Eco*RV-to-*Pst*I (polylinker site) DNA fragment from pLOI682 and lacked PGM activity. Subclone pLOI686 was constructed by deletion of the *Hpa*I-to-*Sst*I (polylinker site) DNA fragment and lacked PGM activity. Subclone pLOI695 was constructed by ligating the *Hpa*I fragment from pLOI682 into the *Hpa*I site of pLOI685 and was active. A slightly shorter subclone was also constructed by a deletion of the *Sty*I-to-*Xba*I (polylinker site) fragment from pLOI682. This clone was also active, effectively localizing the *pgm* coding region to within 871 bp.

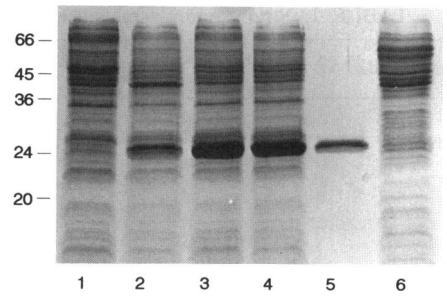

FIG. 2. Sodium dodecyl sulfate-polyacrylamide gel of proteins from recombinant *E. coli*. Positions of molecular weight markers (in thousands) are on the left. Lanes: 1, *E. coli* DH5α(pUC18); 2, *E. coli* DH5α(pLOI680); 3, *E. coli* DH5α(pLOI695); 4, *E. coli* DH5α (pLOI696); 5, purified PGM; 6, *Z. mobilis* CP4.

Southern hybridization analysis. Southern analysis of chromosomal DNA confirmed that the cloned *pgm* gene originated from *Z. mobilis* CP4 (data not shown). The *pgm*-specific probe did not hybridize to *E. coli* DH5α DNA and bound single fragments in genomic digestions with *Eco*RI, *Hin*dIII, *Pst*I, and *Sal*I, indicating that only a single copy resided on the chromosome.

Expression of the *Z. mobilis* gene encoding PGM in *E. coli*. Clones containing *pgm* were analyzed on sodium dodecyl sulfate-polyacrylamide gels (Fig. 2). A prominent protein band which was absent in DH5α(pUC18) (Fig. 2, lane 1) was observed in positive clones (lanes 2 through 4). This band corresponds in size to purified PGM (Fig. 2, lane 5), with an apparent M_r of 26,000. Clones with the most prominent band in this region also exhibited the highest PGM-specific activities.

Additional subclones were constructed in pBluescript II KS+ with the *Z. mobilis* DNA fragment from pLOI680 in the opposite orientation with respect to the *lac* promoter. These clones expressed PGM activity comparable to that of pLOI680, indicating that the *lac* promoter was not required.

A second new protein band with an apparent M_r of 42,000 was observed in the original clone containing pLOI680. By comparing denaturing gels of protein extracts from subclones (data not shown), this gene was localized to the *Z. mobilis* DNA fragment in pLOI683. Expression of this protein in subclones was dependent on the *lac* promoter.

Sequence of clone pLOI680 and identification of ORFs. The full DNA fragment cloned in pLOI680 was sequenced in both directions and contained 5,227 bp (Fig. 3). This sequence contains one partial and four complete ORFs.

The first ORF (ZM1) extends from base 51 through 848 and encodes a 265-amino-acid protein with a predicted M_r of 29,150. This ORF is preceded by a Shine-Dalgarno sequence (GGAG) upstream from TTG, the proposed start codon (34). No significant matches to the translated ZM1 sequence were found in the current data base. No abundant protein was observed on denaturing gels of protein extracts from *E. coli*

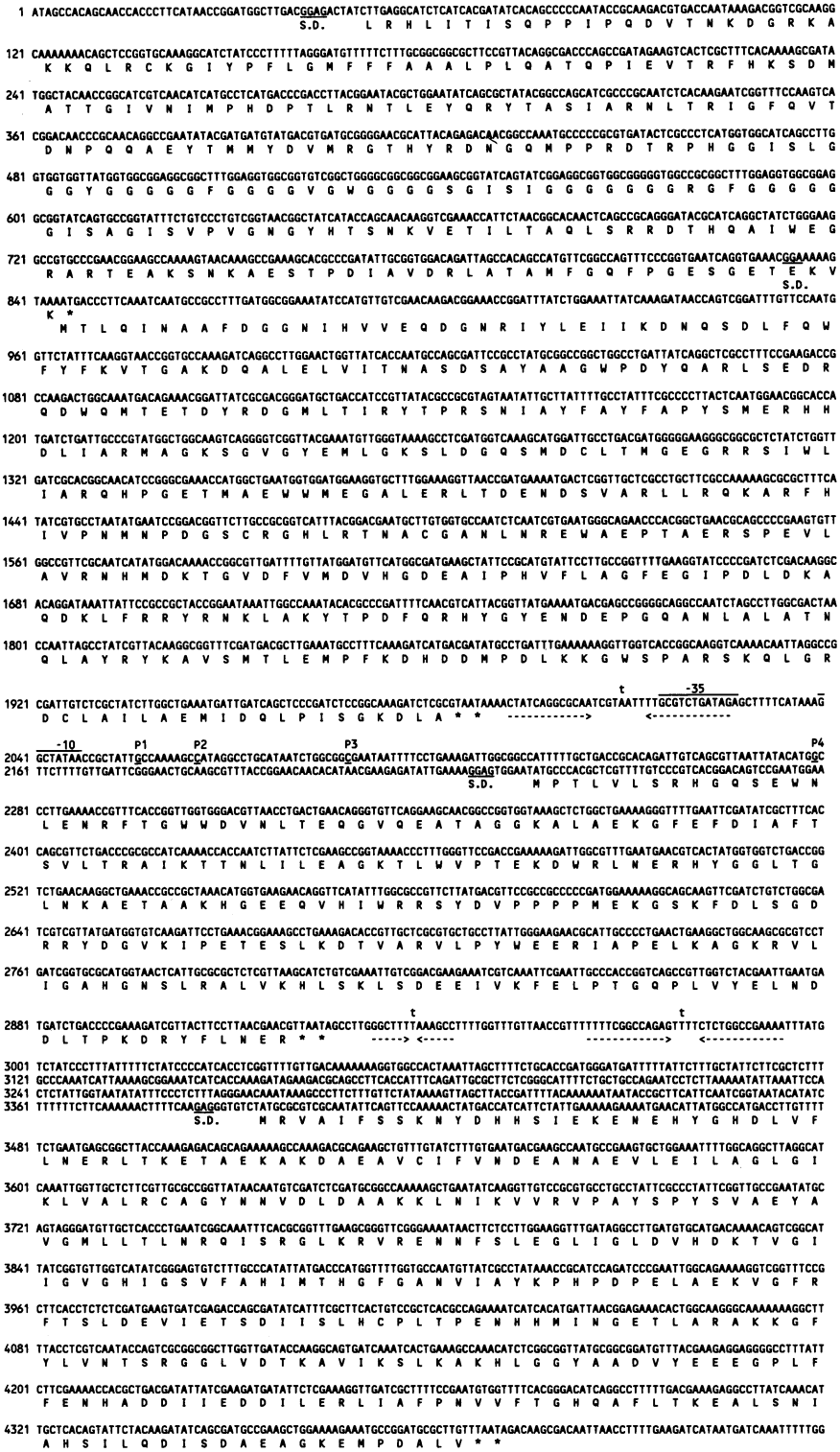

FIG. 3. Nucleotide sequence of the 5.2-kbp DNA fragment from *Z. mobilis* contained in pLOI680. Translated amino acid sequences for ORFs are below the corresponding DNA sequences. Shine-Dalgarno sites (S.D.) are underlined. Stop codons are indicated by asterisks. The four promoters (P1 through P4) for *pgm* are underlined. The −10 and −35 regions of the P1 promoter for *pgm* are overlined. Potential transcriptional terminators (t) are marked with arrows to indicate the palindrome.

subclones which corresponded to this ORF, regardless of orientation with respect to the *lac* promoter. The possibility that ZM1 is an intergenic sequence cannot be excluded.

The second ORF (ZM2) encodes a 380-amino-acid protein (M_r, 41,800) and extends from base 845 through 1,987, overlapping the 3′ end of ZM1. The ATG start codon for ZM2 is preceded by GGA, the proposed ribosome-binding site. This ORF ends with two consecutive stop codons immediately followed by a palindromic sequence which may represent part of a transcriptional terminator. No significant matches to the translated ZM2 sequence were found in the current data base. This protein is expressed in *E. coli* DH5α(pLOI680) as the weaker recombinant protein band on denaturing gels, with an apparent M_r of 42,000 (Fig. 2, lane 2).

The third ORF extends from base 2,237 through 2,923 and encodes *Z. mobilis* PGM (*pgm*), which is 228 amino acids long with an M_r of 25,410. The predicted amino acid sequence for this protein is identical to results from N-terminal protein sequencing for the first 40 amino acids. The coding region of *pgm* is preceded by a ribosome-binding site (GGAG) located upstream from the ATG start codon. This ORF is closed by two consecutive stop codons and immediately followed by two palindromic sequences with A+T-rich regions which may serve as transcriptional terminators.

The fourth ORF, designated *ddh*, extends from base 3,395 through 4,390 and encodes a 331-amino-acid protein (M_r, 36,410). The coding region for *ddh* is preceded by a ribosome-binding site (GAG) upstream from the ATG start codon. This ORF is closed by two consecutive stop codons followed by a palindromic sequence and an A+T-rich region which may serve as a transcriptional terminator. The translated *ddh* sequence is similar to that of D-isomer-specific 2-hydroxyacid dehydrogenases (17). A comparison between the translated amino acid sequences of *ddh* from *Z. mobilis* and D-lactate dehydrogenase from *Lactobacillus plantarum* (35) indicates that these two proteins exhibit 31.5% identity and 55.4% similarity. The translated *Z. mobilis ddh* sequence retains the NADH-binding domain (17) and many of the amino acids which are highly conserved in this class of enzymes (35). No overexpressed protein corresponding to this gene was observed in protein extracts from *E. coli* recombinants.

The fifth ORF was incomplete, but it was readily identified as the N terminus of *Z. mobilis* ADHI (*adhA*), previously sequenced in our laboratory (21).

Mapping of a transcriptional initiation site for *Z. mobilis pgm*. Five potential sites of transcriptional initiation were identified by primer extension analysis in the upstream untranslated region (Fig. 4). Four of these bands are in regions that have partial homology to the consensus sequence for the −10 and −35 upstream regions in *Z. mobilis* (19). The 5′ initiation site is designated P1 (Fig. 4). One band was not included as a putative promoter because it resides at the 3′ border of a prominent stem-loop and because the sequence in the upstream region did not appear to be related to the proposed promoter consensus sequences. It is likely that this band represents a reverse transcriptase pause site.

Considerable similarity was observed in the −10 region, the −35 region, and the connecting sequence, consistent with all being functional. Promoters 3 and 4 appear to be most active on the basis of primer extension analysis.

Codon usage. The patterns of codon usage in ZM2, *pgm*, and *ddh* products were compared with that in the combined *Z. mobilis* glycolytic and nonglycolytic enzymes (data not shown). A single codon is clearly dominant for 17 of the 20 amino acids in the glycolytic enzymes, many of which are favored to a lesser extent among nonglycolytic enzymes. The dominant codon for two amino acids (asparagine and isoleucine) differed for genes encoding glycolytic and nonglycolytic enzymes. Four codons used infrequently (or not at all) by glycolytic enzymes are utilized more often in nonglycolytic enzymes. Codon usages in ZM1 and ZM2 are similar. Codon usages in these genes and in *ddh* resemble that in genes encoding nonglycolytic enzymes. In contrast, codon usage in PGM matched the usage in glycolytic enzymes with two exceptions. The AGC codon was preferred for serine in PGM, while both codons are used equivalently in other glycolytic enzymes. Also, two codons (ATT and ATC) were used almost equally for isoleucine in PGM, whereas ATC is the dominant codon among combined glycolytic enzymes.

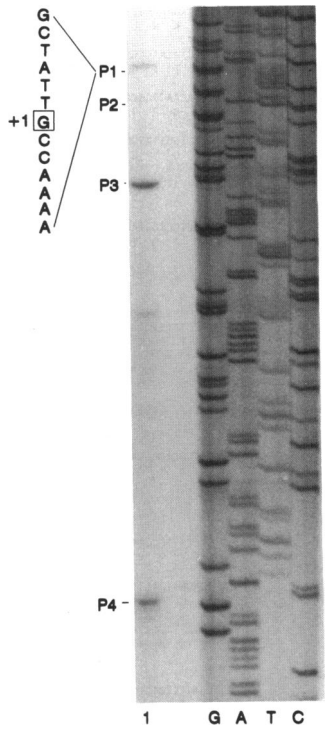

FIG. 4. Primer extension analysis of *pgm* mRNA. Lane 1, putative promoter sites (P1 to P4). Lanes G, A, T, and C refer to the respective dideoxy nucleotide terminators used to generate a sequencing ladder. The expanded sequence for the coding strand of P1 shown is the complement of the sequence ladder. Other promoters are marked on the sequence in Fig. 3.

Conservation of primary amino acid sequence among genes encoding PGMs and BPGA mutases. Comparison of the deduced amino acid sequence of PGM with other sequences revealed 70 to 78% similarity and 51 to 60% identity with BPGA-dependent PGM enzymes from other organisms (Table 2; Fig. 5). The amino acids proposed to be essential for catalysis were also conserved in *Z. mobilis* PGM. A similarly high conservation of primary sequences was observed for the eukaryotic BPGA mutases (Table 2). Despite the high degree of conservation among PGM and BPGA mutases, the *Z. mobilis* PGM was shorter than corresponding enzymes in other organisms and lacked the carboxy terminus and the terminal lysine found in all other PGMs (16). Translations were examined in all three reading frames, and none shared

TABLE 2. Comparison of amino acid sequences for PGMs and BPGA mutases from different organisms

Organism (reference)	% Similarity (% identity)					
	PGM				BPGA mutase	
	S. coelicolor	*S. cerevisiae*	Rat	Human	Human	Mouse
PGM						
Z. mobilis	69.8 (52.0)	70.9 (55.1)	77.4 (60.4)	78.4 (59.0)	70.4 (52.7)	70.2 (51.3)
S. coelicolor (38)		80.0 (64.5)	68.0 (52.8)	67.6 (51.2)	65.0 (48.0)	64.8 (46.0)
S. cerevisiae (36)			67.3 (53.1)	66.9 (51.8)	72.3 (51.4)	72.7 (49.8)
Rat (6)				94.5 (91.7)	71.5 (51.8)	71.5 (50.6)
Human (33)					72.3 (51.4)	72.7 (49.8)
BPGA mutase						
Human (20)						96.9 (92.3)
Mouse (22)						

homology with the eukaryotic carboxy terminus. To further establish the absence of the terminal segment in *Z. mobilis* PGM, two subclones (pLOI696 and pLOI697 [the polymerase chain reaction DNA fragment in pLOI706EH]) which lacked additional sequence beyond the proposed translational stop were constructed. Both expressed large amounts of *Z. mobilis* PGM. Equivalent levels of PGM protein were observed in denaturing gels of DH5α(pLOI696) lacking this 3′ sequence and of DH5α(pLOI695) containing adjacent 3′ sequence (specific activities of 140 and 148 IU/mg of protein, respectively). Together, these results demonstrate that the additional sequence is not essential for the activity of *Z. mobilis* PGM, and they are consistent with the absence of this tail segment.

Effect of overexpression of PGM on growth and glycolytic flux. The rates of growth and carbon dioxide production (a measure of glycolytic flux) were compared in recombinant strains of *Z. mobilis* CP4 containing the vector (pLOI706EH) and additional copies of *pgm* (pLOI697). In this construct, *pgm* is expressed from the *tac* promoter under control of $lacI^q$. Upon induction with 4 mM IPTG, cytoplasmic PGM activity doubled (data not shown). However, the increase in PGM activity did not affect either the specific growth rate (0.44 doublings per h) or the rate of CO_2 evolution (2.1 IU/mg of cell protein).

```
Zym      MPTLVLSRHGQSEWNLENRFTGWWDVNLTEQGVQEATAGGKALAEKGFEFDIAFTSVLTR   60
Str  MADAPYKLILLRHGESEWNEKNLFTGWVDVNLTPKGEKEATRGGELLKDAGLLPDVVHTSVQKR   64
Sac      MPKLVLVRHGQSEWNEKNLFTGWVDVKLSAKGQQEAARAGELLKEKKVYPDVLYTSKLSR   60
           *.* ***.****  * **** **.*  *  **  .*  * .     *.  **   *
                 = =       =                                    =

Zym  AIKTTNLILEAGKTLWVPTEKDWRLNERHYGGLTGLNKAETAAKHGEEQVHIWRRSYDVPPPPM  124
Str  AIRTAQLALEAADRHWIPVHRHWRLNERHYGALQGKDKAQTLAEFGEEQFMLWRRSYDTPPPAL  128
Sac  AIQTANIALEKADRLWIPVNRSWRLNERHYGDLQGKDKAETLKKFGEEKFNTYRRSFDVPPPPI  124
     ** *  . ** .   *.*  . ********* * * .**.*    ***    .***.* ***.

Zym  EKGSKFDLSGDRRYDGVK---IPETESLKDTVARVLPYWEERIAPELKAGKRVLIGAHGNSLRA  185
Str  DRDAEYSQFSDPRYAMLPPELRPQTECLKDVVGRMLPYWFDAIVPDLLTGRTVLVAAHGNSLRA  192
Sac  DASSPFSQKGDERYKYVDPNVLPETESLALVIDRLLPYWQDVIAKDLLSGKTVMIAAHGNSLRG  188
     . .     .* **    *   *.**.*   . * **** . *  .*  *. *...********.
                                                    =

Zym  LVKHLSKLSDEEIVKFELPTGQPLVYELNDDLTP-KDR--YFLNER                    228
Str  LVKHLDGISDADIAGLNIPTGIPLSYELNAEFKPLNPGGTY-LDPDAAAAAIEAVKNQGKKK  253
Sac  LVKHLEGISDADIAKLNIPTGIPLVFELDENLKPSKPS--YYLDPEAAAAGAAAVANQGKK  247
     *****  .** .*  ...*** ** .**. .. *      * *.  ***   ** *** **
```

FIG. 5. Comparison of *Z. mobilis* PGM with homologs from other organisms. Identical residues are indicated by asterisks; conserved residues are indicated by dots; residues considered important for activity are double underlined. To maximize homology, gaps (indicated by dashes) were inserted within sequences. *Zym*, *Z. mobilis* PGM; *Str*, *S. coelicolor* PGM (38); *Sac*, *S. cerevisiae* PGM (36).

DISCUSSION

Structure of the *pgm* gene. The structure of the *pgm* gene is similar to that of the other genes encoding glycolytic and fermentative enzymes in *Z. mobilis* in that it includes most of the features associated with high-level expression (7, 15, 19, 34). The *pgm* gene contains a canonical Shine-Dalgarno region, GGAG, identical to that of *Z. mobilis eda*, *pgi*, *gap*, and *pgk*. This region is spaced optimally, 5 bases upstream from the ATG translational start site. Amino acids in these highly expressed proteins are encoded by a subset of codons. A similar pattern of codon usage, which differs from that for *Z. mobilis* genes expressed at lower levels, is retained for the *Z. mobilis* genes encoding glycolytic enzymes and *pgm*. The *pgm* ORF is followed by two palindromic sequences with A+T-rich regions which may serve as *rho*-independent terminators (8). These secondary structures may also function to block 3′-to-5′ exoribonuclease activity and contribute to mRNA stability (7).

The *pgm* gene contains multiple promoters within a long upstream untranslated region which can fold into a variety of stable structures. Although not examined, it is likely that the *pgm* message, like those of other *Z. mobilis* glycolytic genes (23, 24), is stable to allow synthesis of large amounts of PGM. Long untranslated regions are present on all *Z. mobilis* genes encoding for glycolytic and fermentative enzymes. Untranslated region sequences on highly expressed bacterial genes such as *E. coli ompA* (7, 15) contain an extreme 5′-terminal stem-loop, which has been linked to message stability.

Southern analysis indicated that only one copy of the *pgm* gene is present on the *Z. mobilis* chromosome. This may also indicate that there is no 2,3-BPGA mutase gene present in *Z. mobilis* and that the cloned gene serves both functions. The lack of a measurable effect on the growth rate or the metabolic flux rate associated with increased expression of *pgm* in recombinant *Z. mobilis* CP4 was not entirely unexpected. Since PGM catalyzes an equilibrium reaction, the relative amounts of the 2-PGA and 3-PGA may not change with increased enzymatic activity.

Z. mobilis* glycolytic promoters.** The sequences for the four proposed *pgm* promoters appear highly conserved and are similar to the *E. coli* σ^{32} and σ^{70} promoters (12, 18). Initial attempts to define the promoter sequence for *Z. mobilis* genes encoding glycolytic enzymes identified A**CT*G in the −35 region and TA*TG*A*T** in the −10 region on the basis of three genes (19). Recently Barnell et al. (4) used a computer program to compare the sequences of eight

promoters and found similarity in the region between −100 and +15 bases. Although no clear consensus emerged, TAGA**T was identified in the −10 region of three promoters and was proposed as a conserved element. Conclusive identification of the *pgm* promoter region and those of other genes encoding glycolytic enzymes awaits genetic analysis and in vitro transcription with purified components.

Conservation among PGM amino acid sequences. The amino acid sequence for *Z. mobilis* PGM is shorter than all others reported. All of the active site residues proposed for *Saccharomyces cerevisiae* in PGM have been conserved in the *Z. mobilis* enzyme (Fig. 5). In *S. cerevisiae*, histidine 9 (numbering for *S. cerevisiae* residues) is phosphorylated by BPGA (37). Histidine 182 is also involved in phosphate transfer, and arginine 59 is involved in binding the carboxyl group of the PGAs. Serine 12 and threonine 21 may provide ligands for binding the transferred phosphate group (36). Although there is a remarkable degree of homology, the *Z. mobilis* sequence lacks a carboxy-terminal region which is highly conserved in other organisms. For *S. cerevisiae*, PGM has been studied extensively and the high-resolution crystal structure has been determined (16). The 14 carboxy-terminal amino acid residues could not be seen in the electron density map, presumably because the PGM forms a flexible tail. This region is susceptible to proteolysis in *S. cerevisiae* and rabbit muscle, and its loss abolishes mutase activity (16, 28). The tail has been considered necessary for PGM activity and is thought to exclude water and to help anchor intermediates at the active site (37). It has been proposed that the conserved lysine residue (second to last residue in the *S. cerevisiae* amino acid sequence) partially compensates for the negative charge introduced by the substrates and stabilizes the transition state (16). A similar study with human erythrocyte BPGA mutase (13) also showed that deletion of the carboxy-terminal region resulted in a loss of catalytic function. In *Z. mobilis*, the terminal arginine in PGM may serve to replace the conserved lysine in other organisms, stabilizing transition states and substrate binding.

The *ddh* gene and other ORFs near *pgm*. Although the identities of gene products for ZM1 and ZM2 remain unknown, it is interesting to note the apparent use of TTG as the start codon in ZM1. This codon is also used in *E. coli* in approximately 1% of the genes sequenced (34) but is limited to those which are not highly expressed. Assuming that ZM1 is expressed in *Z. mobilis*, these ORFs are translationally coupled with overlapping coding regions. Both genes may be part of an operon which ends with a *rho*-independent terminator 201 bp upstream from *pgm*.

The *ddh* gene, like *pgm*, appears to be monocistronic. Although the precise function of *ddh* remains unknown, this gene has considerable similarity to all sequenced D-specific 2-hydroxyacid dehydrogenases (17, 35). Interestingly, this family of genes includes those encoding D-lactate dehydrogenase, which uses pyruvate as a substrate (35), and 3-PGA dehydrogenase, which uses the same substrate as PGM (17). Both enzymes utilize substrates that come from glycolysis and in *Z. mobilis* may represent branch points for metabolism. The 3-PGA dehydrogenase is the first enzyme in the biosynthetic pathway to serine and glycine in *E. coli*.

In *Z. mobilis*, the *ddh* appears to form part of a gene cluster with related functions in central metabolism which includes *pgm* upstream and *adhA* downstream. All three are transcribed in the same direction. Other genes for glycolytic enzymes are also grouped together on the *Z. mobilis* genome. The *glf* operon includes the genes encoding *glf*, *zwf*, *edd*, and *glk* (4), and the *gap* operon encodes *gap* and *pgk* (10). In this instance, the two genes *pgm* and *adhA* are not organized into an operon but are separated only by *ddh*, which may have a related function. Such organization can be presumed to have an evolutionary advantage.

ACKNOWLEDGMENTS

This research was supported by a grant from the Department of Energy, Division of Basic Energy Sciences (DE-FG05-86ER1313575).

We thank LI-COR for generously providing the use of a model 4000 DNA sequencer.

REFERENCES

1. **Aldrich, H. C., L. McDowell, M. D. F. S. Barbosa, L. P. Yomano, R. K. Scopes, and L. O. Ingram.** 1992. Immunocytochemical localization of glycolytic and fermentative enzymes in *Zymomonas mobilis*. J. Bacteriol. **174:**4504–4508.
2. **An, H., R. K. Scopes, M. Rodriguez, K. F. Keshav, and L. O. Ingram.** 1991. Gel electrophoretic analysis of *Zymomonas mobilis* glycolytic and fermentative enzymes: identification of alcohol dehydrogenase II as a stress protein. J. Bacteriol. **173:**5975–5982.
3. **Arfman, N., V. Worrell, and L. O. Ingram.** 1992. Use of the *tac* promoter and *lacI*^{q} for the controlled expression of *Zymomonas mobilis* fermentative genes in *Escherichia coli* and *Zymomonas mobilis*. J. Bacteriol. **174:**7370–7378.
4. **Barnell, W. O., J. Liu, T. L. Hesman, M. C. O'Neill, and T. Conway.** 1992. The *Zymomonas mobilis glf, zwf, edd*, and *glk* genes form an operon: localization of the promoter and identification of a conserved sequence in the regulatory region. J. Bacteriol. **174:**2816–2823.
5. **Byun, M. O.-K., J. B. Kaper, and L. O. Ingram.** 1986. Construction of a new vector for the expression of foreign genes in *Zymomonas mobilis*. J. Ind. Microbiol. **1:**9–15.
6. **Castella-Escola, J., L. Montoliu, G. Pons, and P. Puigdomenech.** 1989. Sequence of rat skeletal muscle phosphoglycerate mutase cDNA. Biochem. Biophys. Res. Commun. **165:**1345–1351.
7. **Chen, L.-H., S. A. Emory, A. L. Bricker, P. Bouvet, and J. G. Belasco.** 1991. Structure and function of a bacterial mRNA stabilizer: analysis of the 5′ untranslated region of *ompA* mRNA. J. Bacteriol. **173:**4578–4586.
8. **Cheng, S.-W. C., E. C. Lynch, K. R. Leason, D. L. Court, B. A. Shapiro, and D. I. Friedman.** 1991. Functional importance of sequence in the stem-loop of a transcription terminator. Science **254:**1205–1207.
9. **Conway, T., M. O.-K. Byun, and L. O. Ingram.** 1987. Expression vector for *Zymomonas mobilis*. Appl. Environ. Microbiol. **53:**235–241.
10. **Conway, T., and L. O. Ingram.** 1988. Phosphoglycerate kinase gene from *Zymomonas mobilis*: cloning, sequencing, and localization within the *gap* operon. J. Bacteriol. **170:**1926–1933.
11. **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.
12. **Cowing, D. W., J. C. Bardwell, E. A. Craig, C. Woolford, R. W. Hendrix, and C. A. Gross.** 1985. Consensus sequence for *Escherichia coli* heat shock gene promoters. Proc. Natl. Acad. Sci. USA **82:**2679–2683.
13. **Craescu, C. T., O. Schaad, M. C. Garel, R. Rosa, and S. Edelstein.** 1992. Structural modeling of the human erythrocyte bisphosphoglycerate mutase. Biochimie **74:**519–526.
14. **de Lorenzo, V., M. Herrero, U. Jakubzik, and K. N. Timmis.** 1990. Mini-Tn*5* transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. J. Bacteriol. **172:**6568–6572.
15. **Emory, S. A., P. Bouvet, and J. G. Belasco.** 1992. A 5′-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. Genes Dev. **6:**135–148.
16. **Fothergill-Gilmore, L. A.** 1989. The phosphoglycerate mutases. Adv. Enzymol. Relat. Areas Mol. Biol. **62:**227–313.

17. **Grant, G. A.** 1989. A new family of 2-hydroxyacid dehydrogenases. Biochem. Biophys. Res. Commun. **165:**1371–1374.
18. **Gross, C. A., D. B. Straus, and J. W. Erickson.** 1990. The function and regulation of heat shock proteins in *Escherichia coli*, p. 167–189. *In* R. I. Morimoto, A. Tissieres, and C. Georgopoulos (ed.), Stress proteins in biology and medicine. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Ingram, L. O., C. K. Eddy, K. F. Mackenzie, T. Conway, and F. Alterthum.** 1989. Genetics of *Zymomonas mobilis* and ethanol production. Dev. Ind. Microbiol. **30:**53–69.
20. **Joulin, V., J. Peduzzi, P.-H. Romeo, R. Rosa, C. Valentin, A. Dubart, B. Lapeyre, Y. Blouquit, M.-C. Garel, M. Goossens, J. Rosa, and M. Cohen-Solal.** 1986. Molecular cloning and sequencing of the human erythrocyte 2,3-bisphosphoglycerate mutase cDNA: revised amino acid sequence. EMBO J. **5:**2275–2283.
21. **Keshav, K. F., L. P. Yomano, H. An, and L. O. Ingram.** 1990. Cloning of the *Zymomonas mobilis* structural gene encoding alcohol dehydrogenase I (*adhA*): sequence comparison and expression in *Escherichia coli*. J. Bacteriol. **172:**2491–2497.
22. **LeBoulch, P., V. Joulin, M. C. Garel, J. Rosa, and M. Cohen-Solal.** 1988. Molecular cloning and nucleotide sequence of murine 2,3-bisphosphoglycerate mutase cDNA. Biochem. Biophys. Res. Commun. **156:**874–881.
23. **Liu, J., W. O. Barnell, and T. Conway.** 1992. The polycistronic mRNA of the *Zymomonas mobilis glf-zwf-edd-glk* operon is subject to complex transcript processing. J. Bacteriol. **174:** 2824–2833.
24. **Mejia, J. P., M. E. Burnett, H. An, W. O. Barnell, K. F. Keshav, T. Conway, and L. O. Ingram.** 1992. Coordination of expression of *Zymomonas mobilis* glycolytic and fermentative enzymes: a simple hypothesis based on mRNA stability. J. Bacteriol. **174:** 6438–6443.
25. **Middendorf, L. R., J. C. Bruce, R. C. Bruce, R. D. Eckles, D. L. Grone, S. C. Roemer, G. D. Sloniker, D. L. Steffens, S. L. Sutter, J. A. Brumbaugh, and G. Patonay.** 1992. Continuous, on-line DNA sequencing using a versatile infrared laser scanner/electrophoresis apparatus. Electrophoresis **13:**487–494.
26. **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.
27. **Pawluk, A., R. K. Scopes, and K. Griffiths-Smith.** 1986. Isolation and properties of the glycolytic enzymes from *Zymomonas mobilis*. The five enzymes from glyceraldehyde-3-phosphate dehydrogenase through pyruvate kinase. Biochem. J. **238:**275–281.
28. **Price, N. C., D. Duncan, and J. W. McAllister.** 1985. Inactivation of rabbit muscle phosphoglycerate mutase by limited proteolysis with thermolysin. Biochem. J. **229:**167–171.
29. **Proudfoot, N. J., M. H. M. Shander, J. L. Manley, M. L. Gefter, and T. Maniatis.** 1980. Structure and *in vitro* transcription of human globin genes. Science **209:**1329–1336.
30. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239:**487–491.
31. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
33. **Shanske, S., S. Sakoda, M. A. Hermodson, S. DiMaurom, and E. A. Schon.** 1987. Isolation of a cDNA encoding the muscle-specific subunit of human phosphoglycerate mutase. J. Biol. Chem. **262:**14612–14617.
34. **Stormo, G. D.** 1986. Translation initiation, p. 195–224. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworths, Boston, Mass.
35. **Taguchi, H., and T. Ohta.** 1991. D-Lactate dehydrogenase is a member of the D-isomer-specific 2-hydroxyacid dehydrogenase family. J. Biol. Chem. **266:**12588–12594.
36. **White, M. F., and L. A. Fothergill-Gilmore.** 1988. Sequence of the gene encoding phosphoglycerate mutase from *Saccharomyces cerevisiae*. FEBS Lett. **229:**383–387.
37. **White, M. F., and L. A. Fothergill-Gilmore.** 1990. Mutase versus synthase: the phosphoglycerate mutase family studied by protein engineering. Biochem. Soc. Trans. **18:**257.
38. **White, P. J., J. Nairn, N. C. Price, H. G. Nimmo, J. R. Coggins, and I. S. Hunter.** 1992. Phosphoglycerate mutase from *Streptomyces coelicolor* A3(2): purification and characterization of the enzyme and cloning and sequence analysis of the gene. J. Bacteriol. **174:**434–440.
39. **Winn, S. I., H. C. Watson, R. N. Harkins, and L. A. Fothergill.** 1981. Structure and activity of phosphoglycerate mutase. Philos. Trans. R. Soc. Lond. B **293:**121–130.