JOURNAL OF BACTERIOLOGY, Dec. 1987, p. 5653–5662
0021-9193/87/125653-10$02.00/0

Vol. 169, No. 12

Glyceraldehyde-3-Phosphate Dehydrogenase Gene from *Zymomonas mobilis*: Cloning, Sequencing, and Identification of Promoter Region

T. CONWAY, G. W. SEWELL, AND L. O. INGRAM*

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

Received 28 May 1987/Accepted 18 September 1987

The gene encoding glyceraldehyde-3-phosphate dehydrogenase was isolated from a library of *Zymomonas mobilis* DNA fragments by complementing a deficient strain of *Escherichia coli*. It contained tandem promoters which were recognized by *E. coli* but appeared to function less efficiently than the enteric *lac* promoter in *E. coli*. The open reading frame for this gene encoded 337 amino acids with an aggregate molecular weight of 36,099 (including the N-terminal methionine). The primary amino acid sequence for this gene had considerable functional homology and amino acid identity with other eucaryotic and bacterial genes. Based on this comparison, the *gap* gene from *Z. mobilis* appeared to be most closely related to that of the thermophilic bacteria and to the chloroplast isozymes. Comparison of this gene with other glycolytic enzymes from *Z. mobilis* revealed a conserved pattern of codon bias and several common features of gene structure. A tentative transcriptional consensus sequence is proposed for *Z. mobilis* based on comparison of the five known promoters for three glycolytic enzymes.

Zymomonas mobilis is an obligately fermentative, gram-negative bacterium which cannot grow in the absence of a fermentable carbohydrate (24, 35). In nature, this organism is found in honey and plant saps and is a common contaminant of commercial ethanolic fermentations. All isolates of this genus exhibit the ability to metabolize only a very limited range of carbohydrates, including glucose, fructose, and sucrose. These sugars are metabolized via the enzymes of the Entner-Doudoroff pathway in *Z. mobilis*, with the production of ethanol and carbon dioxide as waste products. The entire energy needs of this organism (homeostasis, biosynthesis, and growth) are dependent on the net production of a single ATP per molecule of hexose consumed (24). Despite an apparent inefficiency of ATP production, this organism is capable of rapid growth, with generation times of 90 to 120 min (28), and exhibits a higher rate of glycolytic flux and ethanol production than *Saccharomyces cerevisiae* (13, 28).

The enzymes of glycolysis are particularly abundant in *Z. mobilis* and are estimated to represent over half of the soluble protein (1). In this organism, substrate-level phosphorylations by pyruvate kinase and glyceraldehyde-3-phosphate dehydrogenase occupy particularly central roles. Previous studies in our laboratory have compared the specific activities of these and other glycolytic enzymes in *Z. mobilis* with the rates of glycolytic flux at different stages of growth and fermentation during batch culture (27). In general, the activities of all glycolytic enzymes appeared to be in excess from in vitro assays under substrate-saturating conditions. The specific activities of all enzymes except pyruvate kinase and glyceraldehyde-3-phosphate dehydrogenase continued to increase after the exponential phase as the rates of glycolytic flux declined. In contrast, the changes in the specific activities of these two enzymes were directly correlated with changes in the rates of glycolytic flux.

Glyceraldehyde-3-phosphate dehydrogenase catalyzes the reversible oxidation and phosphorylation of glyceraldehyde-3-phosphate to produce 1,3-diphosphoglycerate (15). This enzyme (in association with 1,3-diphosphoglycerate kinase) represents a common link between ATP production and the extent of NAD^+ reduction and thus occupies a unique position in glycolysis. In *Z. mobilis* (29) and other organisms (15), this enzyme is composed of four identical subunits. It has been purified to homogeneity from *Z. mobilis* and is reported to represent over 5% of the protein in soluble extracts (29). Similarly high levels of this enzyme have been reported for mammalian muscle tissue and fermenting *S. cerevisiae* (15). The monomer molecular weight of the *Z. mobilis* enzyme has been estimated to be between 32,500 and 41,000 (29); those from other organisms are typically 36,000 (15). The kinetics of glyceraldehyde-3-phosphate dehydrogenases are complex and typically do not follow simple Michaelis-Menten kinetics. This enzyme represents an important model system for investigations of protein evolution, structure, and function (15, 33, 36). The primary structures

TABLE 1. Plasmids and strains used

Strain or plasmid	Relevant genotype	Source or reference
Z. mobilis CP4	Prototroph	28
E. coli		
DF221	*gap*	16
TC4	*recA*	7
DH5α	Δ*lacZ*M15 *recA*	BRL[a]
JM101	Δ*lacZ*M15 F′	BRL
Plasmids		
pLOI193	*cat tet*	7
pLOI310	*gap cat*	This study
pLOI311	*gap bla*	This study
pLOI312	*gap bla*	This study
pLOI314	*gap bla*	This study
pUC18	*bla lacI′Z′*[b]	BRL

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

* Corresponding author.

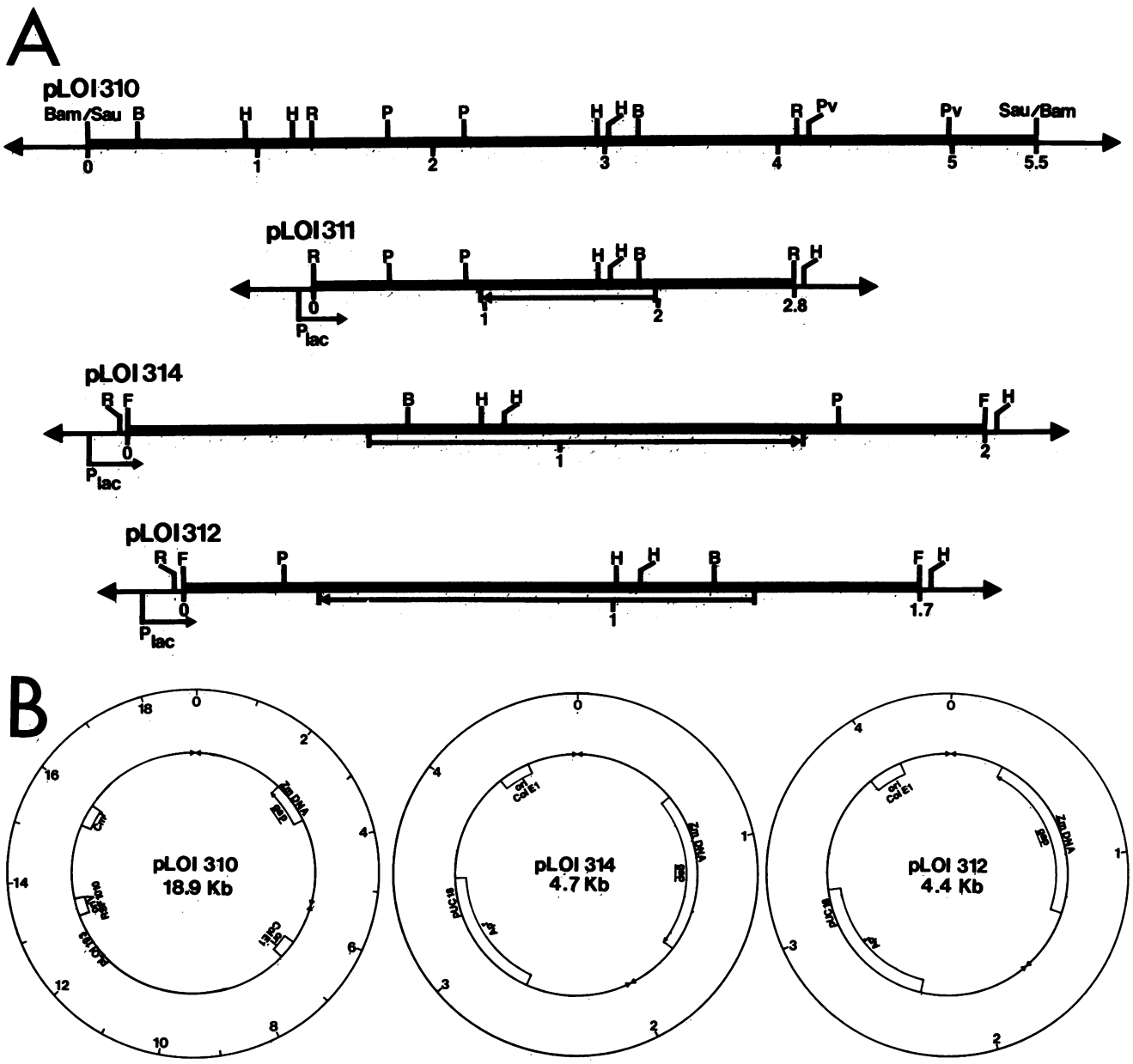


FIG. 1. Restriction maps of DNA fragments (A) and plasmids (B) containing the *gap* gene from *Z. mobilis* (Zm). The bold arrow represents the *gap* gene. No *Bam*HI sites were retained at the junctions of insert and vector DNA in the original construction (pLOI310). The *gap* gene was localized on a 2.8-kb *Eco*RI fragment within pLOI310, and this fragment cloned into the *Eco*RI site of pUC18 to produce pLOI311. Deletions of this 2.8-kb fragment were subloned into the *Sma*I site of pUC18 in both orientations with respect to the *lac* promoter (pLOI314 and pLOI312). Abbreviations: Cm^r, chloramphenicol resistance gene; Ap^r, ampicillin resistance gene; *ori*, origin of replication from either RSF1010 (*oriV*) or ColE1; Bam, *Bam*HI; Sau, *Sau*3A; B, *Bst*EII; R, *Eco*RI; P, *Pst*I; H, *Hin*dIII; Pv, *Pvu*II; F, site of blunt fusion. Numbers under lines denote the DNA size (in kilobases).

and DNA sequences of this enzyme have been determined in many organisms, and all have considerable amino acid homology. Three-dimensional structures of several of these have been established and clearly identify conserved clefts involved in substrate binding, cofactor binding, and subunit interaction (6, 15, 36).

In this study we have cloned and sequenced glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) from *Z. mobilis* as a first step in our investigation of the control of its expression.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The strains and plasmids used in this study are listed in Table 1. *Z. mobilis* CP4 was grown at 30°C in complex medium (28) containing 20 g of glucose per liter. *Escherichia coli* strains TC4 and DH5α were grown in Luria broth (21) or on Luria agar (1.5% agar) lacking added carbohydrate. Strain DF221 (CGSC 5584) was grown in minimal medium as described by Hillman and Fraenkel (16), containing 4 g of sodium suc-

TABLE 2. Specific activity of glyceraldehyde-3-phosphate dehydrogenase clones

E. coli strain	Plasmid	Glyceraldehyde-3-phosphate dehydrogenase sp act (IU/mg)
DF221	None	<0.02
TC4	None	0.55
DF221	pLOI310	0.46
	pLOI312	0.25
	pLOI314	4.53

cinate and 1 g of glycerol per liter. Carbohydrates were sterilized by filtration and added to autoclaved basal medium after cooling. Selections for complementation of the glyceraldehyde-3-phosphate dehydrogenase defect in strain DF221 were made on Luria agar lacking added carbohydrate and on minimal plates containing glucose (16) 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. Strains of *E. coli* were transformed by the method of Mandel and Higa (22).

Cloning the gene encoding glyceraldehyde-3-phosphate dehydrogenase. A library of *Z. mobilis* chromosomal DNA was constructed previously (10) by inserting 5- to 7-kilobase (kb) fragments from a *Sau*3A partial digest into the *Bam*HI site of the tetracycline resistance gene in pLOI193 (7). The resulting plasmids retained chloramphenicol resistance as a selectable marker. Plasmids were prepared from the pooled colonies of the original transformation by the alkaline lysis procedure (23). These plasmid preparations were used as a source of *Z. mobilis* DNA for complementation of the glyceraldehyde-3-phosphate dehydrogenase defect in *E. coli* DF221. Strain DF221 does not grow on Luria agar lacking carbohydrate or on glucose minimal medium (16).

Assay of glyceraldehyde-3-phosphate dehydrogenase. Cells were harvested and disrupted as described previously (10). Enzyme activity was determined by the arsenolysis procedure (5). Activities are expressed as international units per milligram of total cell protein. Protein was determined by the method of Lowry et al. (20) as described by Layne (19), with bovine serum albumin as a standard.

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

Determination of DNA sequence. The *gap* gene from *Z. mobilis* was sequenced by a modification (8) of the dideoxy method of Sanger et al. (32). The entire sequence of the insert contained in pLOI312 was sequenced in both directions by using an overlapping set of Bal31-generated deletions which had been cloned into derivatives of phage M13. Sequence data were analyzed with the programs of Pustell and Kafatos (30), obtained from International Biotechnologies, Inc. (New Haven, Conn.).

Analysis of transcriptional initiation. The 5′ termini of transcripts were mapped by the primer extension method as described previously (9). The primer used for these experiments was synthesized with an Autogen 500 oligonucleotide synthesizer (Millipore/Genetic Design, Bedford, Mass.) and was complementary to the noncoding strand of the *gap* gene (CAACGCTAATTACCG) spanning (3′) base pair (bp) 169 through (5′) bp 183. A sequencing ladder was prepared with the same primer to map the transcriptional starts.

Enzymes and chemicals. Restriction enzymes were obtained from Bethesda Research Laboratories (Gaithersburg, Md.) and from International Biotechnologies, Inc. Reverse transcriptase from avian myeloblastosis virus and other nucleic acid-modifying enzymes were purchased from Bethesda Research Laboratories. Radioactive compounds were purchased from New England Nuclear Corp. (Boston, Mass.). Glyceraldehyde-3-phosphate was purchased from Sigma Chemical Co. (St. Louis, Mo.).

RESULTS AND DISCUSSION

Cloning of the gene encoding glyceraldehyde-3-phosphate dehydrogenase (*gap*). Plasmids containing a library of *Z. mobilis* CP4 DNA were transformed into *E. coli* DF221. Transformants were selected for growth on Luria broth plates containing chloramphenicol. Ten clones were isolated from plates which contained approximately 10,000 chloramphenicol-resistant recombinants (estimated from sample plating on minimal medium containing chloramphenicol, glycerol, and succinate). Based on digestions with *Pvu*II, these fell into three groups and contained 4- to 6-kb inserts. One of these, designated pLOI310, was chosen for further study (Fig. 1).

Plasmids were immediately transferred into strain TC4 or strain DH5α for further study to avoid potential rearrangements and concatenation, which occur in strain DF221. Subclones were constructed in the polylinker region of pUC18 and transformed into strain DH5α by using the insertional inactivation of β-galactosidase activity as a marker. Plasmids were isolated and checked for the presence of a functional *Z. mobilis gap* gene by transformation into strain DF221. Both *Hin*dIII and *Bst*EII were initially identified as being within the *Z. mobilis gap* gene during subcloning. This gene was subsequently localized within a 2.8-kb *Eco*RI fragment (pLOI311). This *Eco*RI fragment was isolated from an agarose gel, treated for increasing times with Bal31 (23), and ligated into the *Sma*I site of pUC18. Representative clones containing inserts of decreasing length were tested for their ability to complement the *gap* defect in strain DF221. Two clones were chosen for further study, pLOI312 and pLOI314. The inserts in plasmids pLOI312 (1.7 kb) and pLOI314 (2.0 kb) are in opposite orientation with respect to the *lac* promoter (Fig. 1).

Expression of glyceraldehyde-3-phosphate dehydrogenase

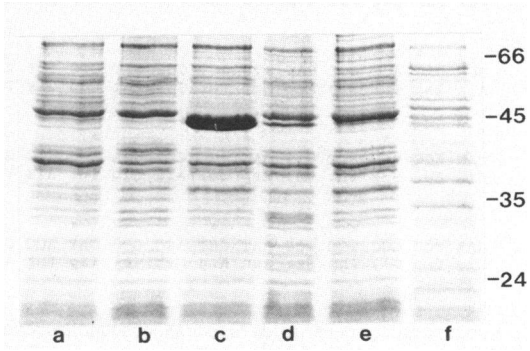


FIG. 2. Sodium dodecyl sulfate-polyacrylamide gel stained with Coomassie blue showing expression of the *Z. mobilis gap* gene in *E. coli*. Lanes: a, strain DF221; b, strain TC4(pUC18); c, strain TC4 (pLOI314); d, strain TC4(pLOI310); e, strain TC4(pLOI312); f, *Z. mobilis* CP4. The positions of molecular weight markers (in thousands) are shown to the right.

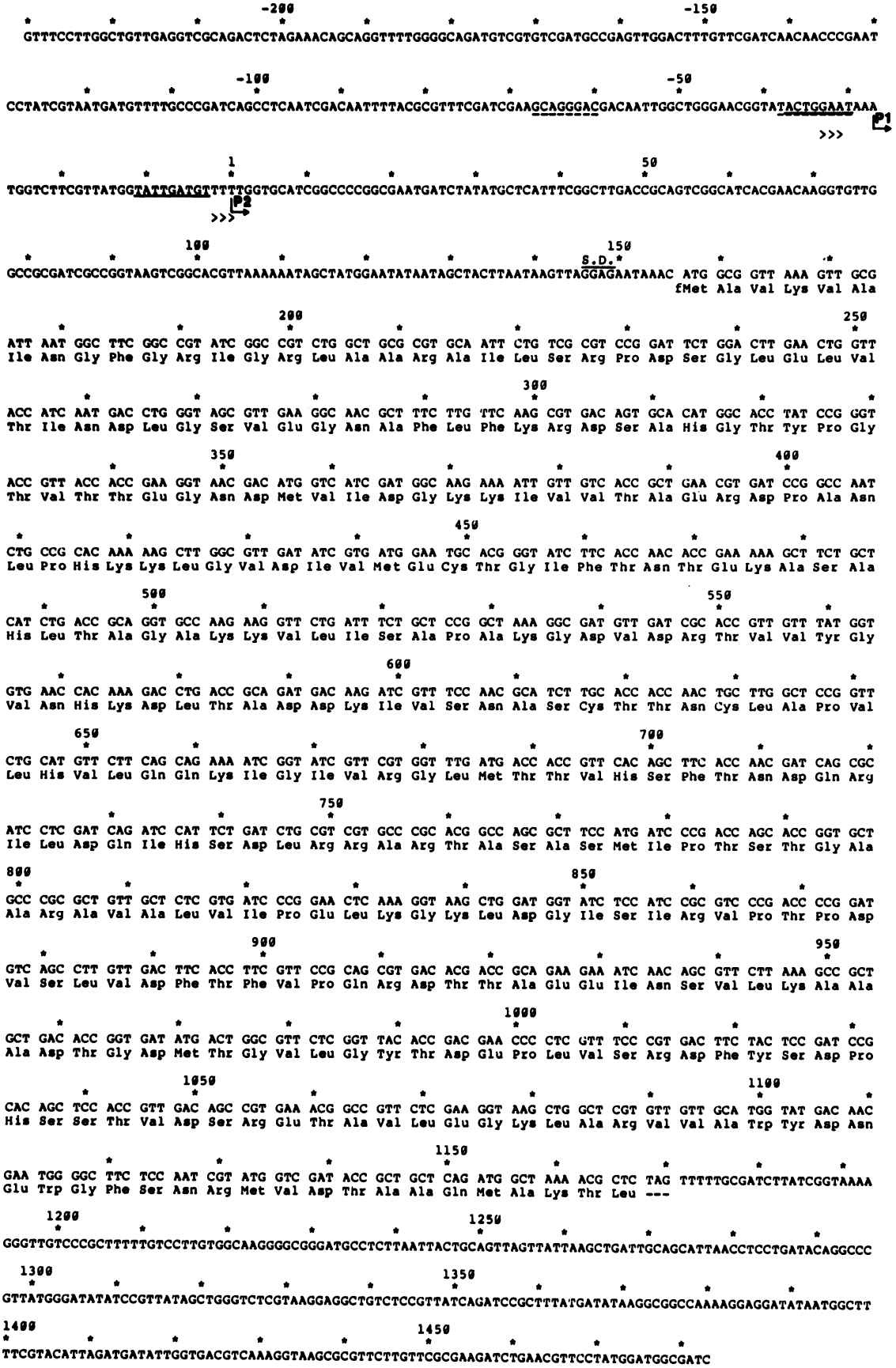

-200
-150
GTTTCCTTGGCTGTTGAGGTCGCAGACTCTAGAAACAGCAGGTTTTGGGGCAGATGTCGTGTCGATGCCGAGTTGGACTTTGTTCGATCAACAACCCGAAT
-100
-50
CCTATCGTAATGATGTTTTGCCCGATCAGCCTCAATCGACAATTTTACGCGTTTCGATCGAAGCAGGGACGACAATTGGCTGGGAACGGTATACTGGAATAAA
P1
>>>
1
50
TGGTCTTCGTTATGGTATTGATGTTTTTGGTGCATCGGCCCCGGCGAATGATCTATATGCTCATTTCGGCTTGACCGCAGTCGGCATCACGAACAAGGTGTTG
P2
>>>
100
150
S.D.
GCCGCGATCGCCGGTAAGTCGGCACGTTAAAAAATAGCTATGGAATATAATAGCTACTTAATAAGTTAGGAGAATAAAC ATG GCG GTT AAA GTT GCG
fMet Ala Val Lys Val Ala
200
250
ATT AAT GGC TTC GGC CGT ATC GGC CGT CTG GCT GCG CGT GCA ATT CTG TCG CGT CCG GAT TCT GGA CTT GAA CTG GTT
Ile Asn Gly Phe Gly Arg Ile Gly Arg Leu Ala Ala Arg Ala Ile Leu Ser Arg Pro Asp Ser Gly Leu Glu Leu Val
300
ACC ATC AAT GAC CTG GGT AGC GTT GAA GGC AAC GCT TTC TTG TTC AAG CGT GAC AGT GCA CAT GGC ACC TAT CCG GGT
Thr Ile Asn Asp Leu Gly Ser Val Glu Gly Asn Ala Phe Leu Phe Lys Arg Asp Ser Ala His Gly Thr Tyr Pro Gly
350
400
ACC GTT ACC ACC GAA GGT AAC GAC ATG GTC ATC GAT GGC AAG AAA ATT GTT GTC ACC GCT GAA CGT GAT CCG GCC AAT
Thr Val Thr Thr Glu Gly Asn Asp Met Val Ile Asp Gly Lys Lys Ile Val Val Thr Ala Glu Arg Asp Pro Ala Asn
450
CTG CCG CAC AAA AAG CTT GGC GTT GAT ATC GTG ATG GAA TGC ACG GGT ATC TTC ACC AAC ACC GAA AAA GCT TCT GCT
Leu Pro His Lys Lys Leu Gly Val Asp Ile Val Met Glu Cys Thr Gly Ile Phe Thr Asn Thr Glu Lys Ala Ser Ala
500
550
CAT CTG ACC GCA GGT GCC AAG AAG GTT CTG ATT TCT GCT CCG GCT AAA GGC GAT GTT GAT CGC ACC GTT GTT TAT GGT
His Leu Thr Ala Gly Ala Lys Lys Val Leu Ile Ser Ala Pro Ala Lys Gly Asp Val Asp Arg Thr Val Val Tyr Gly
600
GTG AAC CAC AAA GAC CTG ACC GCA GAT GAC AAG ATC GTT TCC AAC GCA TCT TGC ACC ACC AAC TGC TTG GCT CCG GTT
Val Asn His Lys Asp Leu Thr Ala Asp Asp Lys Ile Val Ser Asn Ala Ser Cys Thr Thr Asn Cys Leu Ala Pro Val
650
700
CTG CAT GTT CTT CAG CAG AAA ATC GGT ATC GTT CGT GGT TTG ATG ACC ACC GTT CAC AGC TTC ACC AAC GAT CAG CGC
Leu His Val Leu Gln Gln Lys Ile Gly Ile Val Arg Gly Leu Met Thr Thr Val His Ser Phe Thr Asn Asp Gln Arg
750
ATC CTC GAT CAG ATC CAT TCT GAT CTG CGT CGT GCC CGC ACG GCC AGC GCT TCC ATG ATC CCG ACC AGC ACC GGT GCT
Ile Leu Asp Gln Ile His Ser Asp Leu Arg Arg Ala Arg Thr Ala Ser Ala Ser Met Ile Pro Thr Ser Thr Gly Ala
800
850
GCC CGC GCT GTT GCT CTC GTG ATC CCG GAA CTC AAA GGT AAG CTG GAT GGT ATC TCC ATC CGC GTC CCG ACC CCG GAT
Ala Arg Ala Val Ala Leu Val Ile Pro Glu Leu Lys Gly Lys Leu Asp Gly Ile Ser Ile Arg Val Pro Thr Pro Asp
900
950
GTC AGC CTT GTT GAC TTC ACC TTC GTT CCG CAG CGT GAC ACG ACC GCA GAA GAA ATC AAC AGC GTT CTT AAA GCC GCT
Val Ser Leu Val Asp Phe Thr Phe Val Pro Gln Arg Asp Thr Thr Ala Glu Glu Ile Asn Ser Val Leu Lys Ala Ala
1000
GCT GAC ACC GGT GAT ATG ACT GGC GTT CTC GGT TAC ACC GAC GAA CCC CTC GTT TCC CGT GAC TTC TAC TCC GAT CCG
Ala Asp Thr Gly Asp Met Thr Gly Val Leu Gly Tyr Thr Asp Glu Pro Leu Val Ser Arg Asp Phe Tyr Ser Asp Pro
1050
1100
CAC AGC TCC ACC GTT GAC AGC CGT GAA ACG GCC GTT CTC GAA GGT AAG CTG GCT CGT GTT GTT GCA TGG TAT GAC AAC
His Ser Ser Thr Val Asp Ser Arg Glu Thr Ala Val Leu Glu Gly Lys Leu Ala Arg Val Val Ala Trp Tyr Asp Asn
1150
GAA TGG GGC TTC TCC AAT CGT ATG GTC GAT ACC GCT GCT CAG ATG GCT AAA ACG CTC TAG TTTTTGCGATCTTATCGGTAAAA
Glu Trp Gly Phe Ser Asn Arg Met Val Asp Thr Ala Ala Gln Met Ala Lys Thr Leu ---
1200
1250
GGGTTGTCCCGCTTTTTGTCCTTGTGGCAAGGGGCGGGATGCCTCTTAATTACTGCAGTTAGTTATTAAGCTGATTGCAGCATTAACCTCCTGATACAGGCCC
1300
1350
GTTATGGGATATATCCGTTATAGCTGGGTCTCGTAAGGAGGCTGTCTCCGTTATCAGATCCGCTTTATGATATAAGGCGGCCAAAAGGAGGATATAATGGCTT
1400
1450
TTCGTACATTAGATGATATTGGTGACGTCAAAGGTAAGCGCGTTCTTGTTCGCGAAGATCTGAACGTTCCTATGGATGGCGATC

TABLE 3. Comparison of codon usage

Amino acid	Codon[a]	E. coli combined[b,c] (mol%)	Z. mobilis[b]			
			gap	pdc	adhB	Combined[d]
Phe	TTT	1.3	0	0.4	0.8	0.4
	TTC	2.2	2.7	2.9	2.1	2.6
Leu	TTA	0.7	0	0.4	0	0.1
	TTG	0.9	0.9	1.8	1.3	1.3
	CTT	0.8	1.5	1.6	1.8	1.6
	CTC	0.8	2.1	1.8	1.3	1.7
	CTA	0.2	0	0	0	0
	CTG	6.8	3.6	3.4	5.2	4.1
Ile	ATT	2.2	1.2	1.1	1.6	1.3
	ATC	3.7	4.5	3.6	3.1	3.7
	ATA	0.2	0	0	0	0
Met	ATG	2.8	2.4	2.3	4.4	3.0
Val	GTT	2.9	7.7	4.1	5.0	5.6
	GTC	1.2	1.5	3.2	2.3	2.3
	GTA	1.8	0	0	0	0
	GTG	2.2	0.9	0.5	0.5	0.6
Ser	TCT	1.3	1.5	0.5	2.1	1.4
	TCC	1.5	2.1	1.4	1.3	1.6
	TCA	0.4	0	0	1.0	1.3
	TCG	0.6	0.3	0	0.5	0.3
	AGT	0.3	0.3	0.5	0	0.3
	AGC	1.4	2.4	1.4	0.5	1.4
Pro	CCT	0.5	0	0.9	1.0	0.6
	CCC	0.3	0.3	0.4	0.3	0.3
	CCA	0.7	0	0.4	0.8	0.4
	CCG	2.5	3.6	2.9	2.1	2.9
Thr	ACT	1.1	0.3	0.7	1.0	0.7
	ACC	2.4	7.4	2.9	2.9	4.4
	ACA	0.3	0	0	0	0
	ACG	0.8	1.5	1.1	1.3	1.3
Ala	GCT	2.6	5.3	7.7	8.9	7.3
	GCC	2.2	2.1	3.0	3.1	2.7
	GCA	2.3	2.1	3.2	2.1	2.5
	GCG	3.2	0.9	1.4	0.5	0.9
Tyr	TAT	1.0	0.9	2.7	1.6	1.7
	TAC	1.5	0.6	1.3	0.8	0.9
His	CAT	0.7	1.2	0.9	1.0	1.0
	CAC	1.2	1.2	1.3	1.3	1.3
Gln	CAA	1.0	0	0.2	0.3	0.2
	CAG	3.2	1.8	2.1	1.6	1.8
Asn	AAT	1.0	1.2	1.1	1.8	1.4
	AAC	2.8	2.7	4.5	3.7	3.6
Lys	AAA	4.1	3.0	3.8	3.7	3.5
	AAT	1.3	2.4	2.7	2.9	2.7
Asp	GAT	2.5	4.5	2.0	3.9	3.5
	GAC	3.0	3.6	2.5	2.3	2.8
Glu	GAA	4.9	3.9	6.1	4.2	4.7
	GAG	1.8	0	0.4	0.3	0.2
Cys	TGT	0.4	0	0.2	0.3	0.2
	TGC	0.5	0.9	1.1	1.3	1.1
Trp	TGG	0.7	0.6	1.1	0	0.6
Arg	CGT	3.1	4.2	1.3	1.6	2.4
	CGC	2.0	1.5	1.6	0.3	1.1
	CGA	0.2	0	0	0	0
	CGG	0.2	0	0.2	0	0.1
	AGA	0.06	0	0	0	0
	AGG	0.04	0	0	0	0
Gly	GGT	3.8	4.2	5.7	6.0	5.3
	GGC	3.1	3.0	2.0	2.3	2.4
	GGA	0.4	0.3	0.2	0	0.2
	GGG	0.6	0	0	0	0

[a] Preferred codons are underlined.
[b] Expressed as a molar percentage of codon usage.
[c] Combination of 52 proteins in *E. coli* (2).
[d] Combination of three glycolytic enzymes from *Z. mobilis*.

gene from *Z. mobilis* in *E. coli*. Table 2 summarizes the specific activity of glyceraldehyde-3-phosphate dehydrogenase in strains of *E. coli* with and without plasmids containing the *gap* gene from *Z. mobilis*. Strain DF221 exhibited very low levels of glyceraldehyde-3-phosphate dehydrogenase activity, as shown previously (16). Activity was readily detected in strain TC4, which carries the wild-type gene. The level of glyceraldehyde-3-phosphate dehydrogenase activity in strain DF221 containing pLOI310 (the original 5.5-kb fragment from *Z. mobilis*) was equivalent to that of wild-type *E. coli*, consistent with the observed complementation of the *gap* gene mutation. Tenfold-higher activity was observed in strain DF221 containing pLOI314, in which the *lac* promoter of pUC18 was oriented in the direction of *gap* transcription. Only 1/20 of this activity was observed in a similar subclone (pLOI312) which was oriented in the opposite direction with respect to the pUC18 *lac* promoter. These results provide evidence that the native promoter for *gap* from *Z. mobilis* is considerably less efficient in *E. coli* than the *lac* promoter. The specific activity of glyceraldehyde-3-phosphate dehydrogenase in *Z. mobilis* ranges from 0.8 to 4 IU/mg of total cell protein (27). *E. coli* DF221 containing pLOI314 contained even higher activities.

Figure 2 shows a sodium dodecyl sulfate-polyacrylamide gel of whole cell proteins from *Z. mobilis* and *E. coli*. Pawluk et al. (29) have reported that the subunit molecular weight of *Z. mobilis* glyceraldehyde-3-phosphate dehydrogenase was 36,000, based on gel filtration of native enzyme, although the denatured subunits migrated in the region corresponding to 41,000 on sodium dodecyl sulfate-polyacrylamide gels. A prominent band in the region corresponding to a molecular weight of 41,000 was clearly visible in Fig. 2 (lane c, strain TC4 containing pLOI314). This strain contained the highest activity of *Z. mobilis* glyceraldehyde-3-phosphate dehydrogenase. Much less prominent bands were also present in this region in strain DF221 lacking genes from *Z. mobilis* (lane a) and in *E. coli* TC4 containing pUC18 (lane b). This band was less prominent in strain TC4 containing pLOI310 (lane d) and

FIG. 3. Sequence and translation of the glyceraldehyde-3-phosphate dehydrogenase gene from *Z. mobilis*. The proposed Shine-Dalgarno (S.D.) sequence is overlined and labeled. Two promoters were found. The promoter nearest the Shine-Dalgarno sequence is labeled P2. Regions corresponding to the −10 and −35 sites are underlined. The distal promoter is labeled P1. Transcriptional initiations by *Z. mobilis* are indicated by arrows. Transcriptional initiations by *E. coli* are indicated by the symbol >.

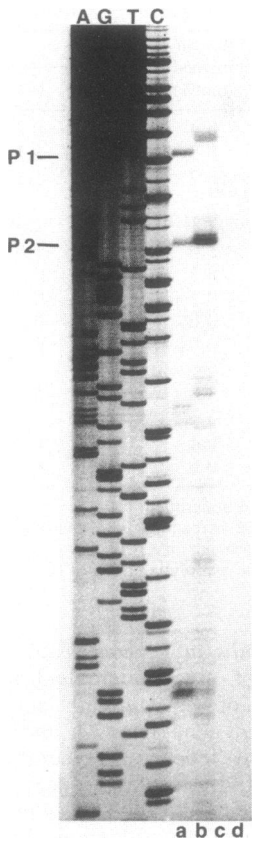


FIG. 4. Primer extension analysis of transcriptional initiation. Lanes: a, RNA from *Z. mobilis* CP4; b, RNA from *E. coli* TC4 containing pLOI312; c, RNA from *E. coli* TC4 containing pLOI314; d, RNA from strain TC4 lacking plasmid. Lanes at the left labeled A, G, T, and C represent a sequencing ladder of the coding strand. The positions of the two promoters recognized by *Z. mobilis* are marked as P1 and P2.

in strain TC4 containing pLOI312 (lane e) than in strain TC4(pLOI314). This putative glyceraldehyde-3-phosphate dehydrogenase band in *E. coli* corresponded to one of the major protein bands in *Z. mobilis* (lane f). From the reported specific activity for purified glyceraldehyde-3-phosphate dehydrogenase, 205 IU/mg (29), this enzyme can be calculated to represent 0.1 to 0.2% of the total cell protein in *Z. mobilis* CP4 and *E. coli* DF221 containing either pLOI310 or pLOI312. Similarly, strain DF221 containing pLOI314 can be estimated to contain over 2% of its total cell protein as this enzyme. Comparison of these calculated values with the gel profiles suggests that the calculated values may underestimate the levels of this *Z. mobilis* protein in recombinant strains of *E. coli*.

Sequence of the glyceraldehyde-3-phosphate dehydrogenase gene. Figure 3 shows a summary of the nucleotide and translated amino acid sequence for the *gap* gene from *Z. mobilis*. This gene contained an open reading frame encoding 337 amino acids (including the N-terminal methionine) with an aggregate molecular weight of 36,099. Table 3 shows a summary of codon usage for *gap*. Seventeen amino acid codons were unused. A proposed ribosome-binding site (AGGAG) occurred 7 bases upstream from the first in-frame ATG. The short intervening region between the ribosome-binding site and the ATG and the 42 bp immediately upstream from the proposed ribosome-binding site were particularly rich in adenine and thymidine. The open reading frame was followed by tandem, in-frame termination codons which were separated by 18 bp. Two inverted repeats which were separated by 3 bp (TTTTTGCGAT/CTT/ATCG GTAAAA) occurred immediately following the first translational terminator. This sequence was followed by a second potential hairpin (GGG/TTGT/CCC). Several sequences were present within the region 200 bp upstream from the proposed ribosome-binding site which had some homology with the proposed enteric bacterial consensus sequence (30) for the −10 and −35 regions. The G+C content of the 1,713 base pairs sequenced was 50.6%, higher than the reported 48.5% G+C composition of bulk genomic DNA from *Z. mobilis* (24). That of the coding region alone was 53.9%.

Transcriptional initiation of *gap* in *E. coli* and *Z. mobilis*. The sites of transcriptional initiation were identified by primer extension with mRNA as a template (Fig. 4). At least two and possibly three transcriptional starts appeared to be present in mRNA from *Z. mobilis*. The two most likely starts occurred over 150 bp upstream from the proposed ribosome-binding site, with one initiation at a thymidine and the other at an adenosine. The proximal promoter (labeled P2) has been assigned as bp +1, with the upstream promoter as base pair −27 (labeled P1). The −35 region of the proximal promoter overlapped the −10 region of the distal promoter (labeled P2). A third band was observed (bp +121) which may represent a third transcriptional initiation or a partial degradation product.

mRNA isolated from *E. coli* TC4(pLOI312) (Fig. 4, lane b) appeared to have 5′ termini mapping near the regions identified as the two main promoter sites for *Z. mobilis*. In both cases, adjacent multiple initiations were observed immediately upstream from the regions identified in *Z. mobilis*. Such multiple initiations may reflect imprecise binding and initiation by *E. coli* RNA polymerase compared with that of *Z. mobilis*. No initiation sites were observed in the region of the *Z. mobilis* promoter by using mRNA from strain TC4(pLOI314) (Fig. 4, lane c), the construction with the highest glyceraldehyde-3-phosphate dehydrogenase activity. In this strain, the only band present was at the top of the gel (not shown) and was presumed to be derived from the *lac* promoter. This very active upstream promoter may have

TABLE 4. Amino acid homology among glyceraldehyde-3-phosphate dehydrogenases from *Z. mobilis* and from other organisms

Sequence[a]	% Identity[b]			
	Cofactor region (aa1–152)	Catalytic region (aa153–336)	S-loop region (aa182–204)	Combined
Z. mobilis	100	100	100	100
E. coli	48.7	47.3	39.1	47.9
B. stearothermophilus	51.3	58.7	60.9	55.4
T. aquaticus	46.0	52.2	60.9	49.4
S. cerevisiae (*tdh1*)	49.3	51.6	39.1	50.6
Homo sapiens (muscle)	44.1	44.6	26.1	44.3
Nicotiana tabacum chloroplast				
gapA	48.0	52.2	60.9	
gapB	47.4	55.4	60.9	

[a] References for sequences are found in the legend to Fig. 5, except for *N. tabacum* (33).
[b] aa, Amino acid residues.

blocked the weaker initiations associated with the *Z. mobilis* promoters. Similar promoter occlusion has been reported for other systems (4, 31).

Comparison of codon usage between the three glycolytic genes of *Z. mobilis* and the aggregate codon usage reported for *E. coli*. Only 39 of the possible 61 amino acid codons were used more than once by *Z. mobilis* in the *gap* gene. Eight codons were not used in *gap* or in the two other *Z. mobilis* genes (CTA, ATA, GTA, ACA, CGA, AGA, AGG, and GGG). All of these codons except GTA are minor codons in *E. coli* (12). A single codon was clearly dominant (underlined in Table 3) for each amino acid except serine, histidine, lysine, and aspartic acid in the *Z. mobilis* genes. None of the dominant codons in *gap*, *adhB*, or *pdc* represent minor codons in *E. coli*. Both *E. coli* and *Z. mobilis* utilized the same codons most frequently for each amino acid except in two cases, alanine and tyrosine. In general, the extent of codon bias in *gap* and the other two genes from *Z. mobilis* was more severe than for the combined genes of *E. coli* (2). Such codon preference may represent bias for abundant tRNAs in *Z. mobilis*, facilitating the biosynthesis (12, 14) of high levels of these three glycolytic enzymes.

Conservation of primary amino acid sequence among glyceraldehyde-3-phosphate dehydrogenases. The primary amino acid sequences of glyceraldehyde-3-phosphate dehydrogenase from all organisms examined show considerable homology (3, 6, 15), as does the sequence from the *Z. mobilis* gene. Figure 5 shows a comparison of the primary sequences between *Z. mobilis* and selected other organisms: *E. coli*, *Bacillus stearothermophilus*, *Thermus aquaticus*, *S. cerevisiae* (*tdh1*), and human muscle. All of these contained large regions of functional amino acid homology (indicated by boxing) and amino acid identity (indicated by shading). Table 4 summarizes the extent of amino acid identity between the *Z. mobilis* enzyme and the five other sequences. Overall, the *Z. mobilis* primary sequence showed the highest homology with that of the *B. stearothermophilus* enzyme (55.4%). Amino acid positions (*Z. mobilis* numbers) which have been identified as being important in binding and catalysis (15), such as cysteine-153, glycine-8, aspartate-35, serine-152, threonine-154, and histidine-180, were all conserved in *Z. mobilis*. The lysine-187 found in most eucaryotic sequences and in *E. coli* (3) has been converted to an arginine in the *Z. mobilis* sequence, analogous to *B. stearothermophilus* (36) and *T. aquaticus* (17).

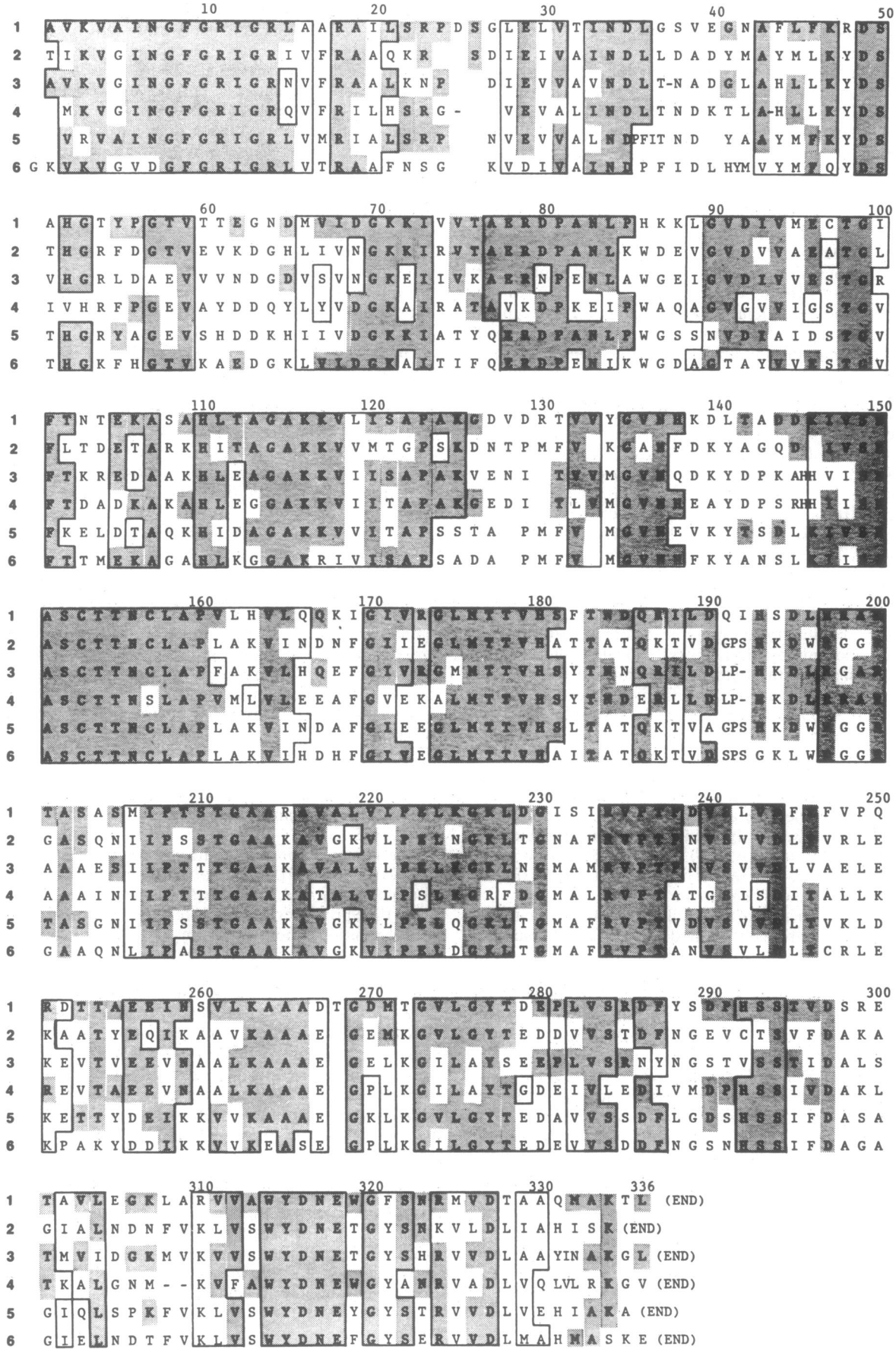
The peptide sequence of glyceraldehyde-3-phosphate dehydrogenase can be roughly divided (15) into a nicotinamide-binding region, catalytic region, and S-loop region (within the catalytic region). This S-loop region is particularly important in that it contains the hydrophobic sequences which form the core of the tetrameric enzyme and important ionic contact points between subunits (26). All genes examined showed a similar level of homology with *Z. mobilis* in the catalytic and nicotinamide-binding regions, except for the catalytic region of *B. stearothermophilus*, in which a somewhat higher level of homology (58.7%) was observed. The S-loop regions of *E. coli*, *S. cerevisiae*, and human muscle showed less homology with this region of the *Z. mobilis* gene. In contrast, the S-loop region of the *T. aquaticus* and *B. stearothermophilus* enzymes exhibited 60.9% homology with that of *Z. mobilis*. Based on this comparison of primary structure, the *Z. mobilis* gene for glyceraldehyde-3-phosphate dehydrogenase appears to be more closely related to that of the two thermophilic bacteria than to either that of *E. coli* or the two eucaryotic sequences examined. It is interesting that the chloroplastic glyceraldehyde-3-phosphate dehydrogenase isozymes also had more homology with the enzymes from the two thermophilic bacteria and the chloroplast enzymes than with eucaryotic enzymes (cytoplasmic) or those of mesophilic bacteria.

The S-loop region has been proposed as being particularly important for thermal stability (26). Despite the apparent similarity of peptide sequence between *Z. mobilis* and the thermophilic bacteria in this region, crude preparations of this enzyme from *Z. mobilis* and from recombinants of *E. coli* DF221(pLOI314) appeared to be more sensitive to thermal inactivation than the native *E. coli* enzyme. After 5 min of incubation at 60°C followed by immediate cooling to 0°C, 51% of the original activity was retained by the native *E. coli* enzyme from strain TC4. The native *Z. mobilis* enzyme from strain CP4 retained only 20% of the original activity after this treatment, and 8% of the original activity was retained by *E. coli* DF221 containing pLOI314.

Comparison of transcriptional and translational control sequences between the three glycolytic genes from *Z. mobilis* and other bacterial sequences. The three glycolytic genes from *Z. mobilis* had several common features. Two of these had tandem promoters, *gap* and *adhB*. All three contained long, untranslated 5′ leader sequences. Although the function of these leader sequences is unknown, such sequences could protect mRNA-coding regions to some extent against degradation.

The Shine-Dalgarno sequences in the glycolytic genes of *Z. mobilis* were analogous to those proposed as being optimal for *E. coli* (34). These were AGGAG in *gap*, GGAG in *pdc*, and GAGGT in *adhB*. The spacing between the ATG and these ribosome-binding sequences was 6 to 7 bases, similar to the average for *E. coli* of 7 ± 1 base (34). In all three genes, intervening sequences between the Shine-Dalgarno and ATG were A+T rich, a composition which has been reported to enhance translation rates in *E. coli* (34). An alanine codon (GCT) immediately following the first in-frame ATG has been reported to be present in highly expressed genes in *E. coli* (34). An alanine codon (GCG or GCT) was present in this position in *gap* and *adhB*. Thus, the translational control sequences on the *Z. mobilis* glycolytic enzymes contain many of the features of highly expressed genes in *E. coli*.

The sequences present in the promoter regions of the three glycolytic genes from *Z. mobilis* also had some degree of homology (Fig. 6A). Although the sample size was quite small (five promoters), several features may prove important for recognition. The −10 regions of the *gap* P2 and *adhB* P2 promoters were remarkably similar in sequence, with identity between 7 of 9 consecutive bases. The −10 sequence of *gap* P1 had more homology with that of the *pdc* promoter than with the *gap* P2 promoter, with identity between 5 of 9 consecutive bases. The −10 region of the *adhB* P1 promoter was most similar to that of the *gap* P1 promoter, with identity between 5 of 9 consecutive bases. A tentative consensus can be deduced in which a position is defined by identity within 4 of the 5 promoters, 5′-TA**G***T-3′. The position of the conserved 5′ T ranged from −10 for *adhB* P2 to −13 for *pdc* with respect to the initiation of transcription (bp +1). Since our assignment of transcriptional initiation by primer extension is generally accurate to within 2 bases, this discrepancy in spacing may well be within experimental error. The average spacing between the −10 region and the translational initiation site was 10.6 bp for the three glycolytic enzymes of *Z. mobilis*. The sequence in the −10 region resembled the enteric bacterial consensus sequence (31) in that it began with the highly conserved TA and ended with


10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
336
(END)

A

-10 Region		mRNA init.	Gene/Promoter
---- G T A T T G A T G T	(2)	T T G >>>	gap P2
---- A T A T T G C T T T	(0)	T G C >>>	adhB P2
---- A T A C T G G A A T	(2)	A T G >>>	gap P1
---- T T A G A G G T T T	(3)	G G T >>>	pdc P
---- G T A G A A A A A T	(1)	T C G >>>	adhB P1
T A T T G * T y T			gap P2 + adhB P2
T A * * G G * * T			gap P1 + pdc P
T A * * * * A A T			gap P1 + adhB P1
T A * * G * * * T			Possible ZM consensus
T A t a a T			Enteric consensus

B

-35 Region	Intervening	Gene/Promoter
--- A A G C A G G G A C G A C ---	21 bp	gap P1
--- A T A C T G G A A T A A A ---	18 bp	gap P2
--- A G C C T T G C T C A T C ---	17 bp	adhB P1
--- C C C C T T G A T C T G A ---	18 bp	adhB P2
--- T G C C T A T A G C T A A ---	19 bp	pdc P
p C * G G p A y p A		gap P1 + P2
C C T T G * T C		adhB P1 + P2
C C T * * * * C		adhB + pdc
* C T * G * * C		Possible Zm consensus
T T G a c a t (17 bp)		Enteric consensus

FIG. 6. Comparison of sequences in promoter regions among the three glycolytic enzymes of *Z. mobilis* (Zm). (A) −10 region; (B) −35 region. Numbers in parentheses in panel A refer to additional bases which separate the sequence in the −10 region from the base position at which mRNA synthesis appears to be initiated (underlined). Dashes denote additional flanking sequences which are untranscribed, and >>> represents additional sequences which are part of the mRNA. A, T, G, and C, base sequence or most highly conserved bases in promoter comparisons; a, t, g, and c, bases which are usually conserved; *, variable bases; y, pyrimidine; p, purine.

the highly conserved T, although the *Z. mobilis* sequence appeared to be longer.

We have tentatively identified a −35 region in the promoter regions of the *Z. mobilis* glycolytic genes (Fig. 6B). The −35 regions of the two *gap* promoters were homologous in 5 of 9 consecutive bases. The −35 regions of the two *adhB* promoters were homologous in 7 of 8 consecutive bases. The −35 region of *pdc* most closely resembled that of the *adhB* gene, sharing 4 of 8 consecutive bases. A tentative consensus sequence can be deduced in which a position is assigned by identity in 4 of the 5 promoters, *CT*G**C. The number of intervening bases between the 3′ conserved C of the −35 region and the 5′ conserved T of the −10 region ranged from 17 to 21 bp, with an average of 18.6 bp. This sequence in the −35 region had some homology with the consensus sequence proposed for *E. coli* (31). The differences between the *Z. mobilis* promoter sequences and the consensus sequence for *E. coli* (31) in the −10 and −35 regions of *gap* are consistent with our observations that these native sequences were less efficient than the enteric bacterial *lac* promoter on derivatives of pUC18. It is interesting that the −35 regions within the *gap* and *adhB* genes which contained tandem promoters were more similar to each other than to those of other genes. In contrast, the −10 regions were more similar between the two distal promoters and between the two proximal promoters of *adhB* and *gap* than they were within a single gene.

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 W. E. Gurley for their suggestions during the preparation of the 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 (no. 8179).

LITERATURE CITED

1. **Algar, E. M., and R. K. Scopes.** 1985. Studies on cell-free metabolism: ethanol production by extracts of *Zymomonas mobilis*. J. Biotechnol. **2**:275–287.
2. **Allf-Steinberger, C.** 1984. Evidence for coding pattern on the non-coding strand of the *E. coli* genome. Nucleic Acids Res. **12**:2235–2241.
3. **Branlant, G., and C. Branlant.** 1985. Nucleotide sequence of the *Escherichia coli gap* gene. Different evolutionary behavior of the NAD^+-binding domain and the catalytic domain of D-glyceraldehyde-3-phosphate dehydrogenase. Eur. J. Biochem. **150**:61–66.
4. **Busby, S., H. Aiba, and B. DeCrombrugghe.** 1982. Mutations in the *Escherichia coli* operon that define two promoters and the binding site of the cyclic AMP receptor protein. J. Mol. Biol. **154**:211–227.
5. **Byers, L. D.** 1982. Glyceraldehyde-3-phosphate dehydrogenase from yeast. Methods Enzymol. **89**:326–335.
6. **Chen, S., T. D. Lee, K. Legesse, and J. E. Shively.** 1986. Identification of the arylazido-β-alanyl-NAD^+-modified site in rabbit muscle glyceraldehyde-3-phosphate dehydrogenase by microsequencing and fast atom bombardment mass spectrometry. Biochemistry **25**:5392–5395.
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, 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.

FIG. 5. Comparison of amino acid homology between selected glyceraldehyde-3-phosphate dehydrogenases from different organisms. Major regions of conserved functional homology are boxed, and amino acids with identity are shaded. Lines: 1, *Z. mobilis*; 2, *E. coli* (3); 3, *B. stearothermophilus* (36); 4, *Thermus aquaticus* (17); 5, *S. cerevisiae* (*tdh1*) (18); 5, *Homo sapiens* muscle (25).

11. **Dawes, E. A., and D. W. Ribbons.** 1966. Sucrose utilization by *Zymomonas mobilis*: formation of a levan. Biochem. J. **98**:804–812.
12. **de Boer, H. A., and R. A. Kastelein.** 1986. Biased codon usage: an exploration of its role in optimization of translation, p. 225–285. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, Stoneham, Mass.
13. **Dombek, K. M., and L. O. Ingram.** 1986. Magnesium limitation and its role in the apparent toxicity of ethanol during yeast fermentation. Appl. Environ. Microbiol. **52**:975–981.
14. **Grosjean, H., and W. Fiers.** 1982. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Genes **18**:199–209.
15. **Harris, J. I., and M. Waters.** 1976. Glyceraldehyde-3-phosphate dehydrogenase, p. 1–49. *In* P. D. Boyer (ed.), The enzymes, vol. XIII. Academic Press, Inc., New York.
16. **Hillman, J. D., and D. G. Fraenkel.** 1975. Glyceraldehyde-3-phosphate dehydrogenase mutants of *Escherichia coli*. J. Bacteriol. **122**:1175–1179.
17. **Hocking, J. D., and J. I. Harris.** 1980. D-Glyceraldehyde-3-phosphate dehydrogenase. Amino acid sequence of the enzyme from the extreme thermophile, *Thermus aquaticus*. Eur. J. Biochem. **108**:567–579.
18. **Holland, M. J., and J. P. Holland.** 1979. The primary structure of a glyceraldehyde-3-phosphate dehydrogenase gene from *Saccharomyces cerevisiae*. J. Biol. Chem. **254**:9839–9845.
19. **Layne, E.** 1957. Spectrophotometric and turbidometric methods for measuring proteins. Methods Enzymol. **3**:447–454.
20. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193**:265–275.
21. **Luria, S. E., and M. Delbruck.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28**:491–511.
22. **Mandel, M., and A. Higa.** 1970. Calcium dependent bacteriophage DNA infection. J. Mol. Biol. **53**:159–162.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **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., Menlow Park, Calif.
25. **Nowak, K., M. Wolny, and T. Banas.** 1981. The complete amino acid sequence of human muscle glyceraldehyde-3-phosphate dehydrogenase. FEBS Lett. **134**:143–146.
26. **Olsen, K. W.** 1983. Structural basis for the thermal stability of glyceraldehyde-3-phosphate dehydrogenases. Int. J. Peptide Protein Res. **22**:469–475.
27. **Osman, Y. A., T. Conway, S. J. Bonnetti, and L. O. Ingram.** 1987. Glycolytic flux in *Zymomonas mobilis*: enzyme and metabolite levels during batch fermentation. J. Bacteriol. **169**: 3726–3736.
28. **Osman, Y. A., and L. O. Ingram.** 1985. Mechanism of ethanol inhibition of fermentation in *Zymomonas mobilis* strain CP4. J. Bacteriol. **164**:173–180.
29. **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.
30. **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.
31. **Reznikoff, W. S., and W. R. McClure.** 1986. *E. coli* promoters, p. 1–33. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, Stoneham, Mass.
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. **Shih, M.-C., G. Lazar, and H. M. Goodman.** 1986. Evidence in favor of the symbiotic origin of chloroplasts: primary structure and evolution of tabacco glyceraldehyde-3-phosphate dehydrogenases. Cell **47**:73–80.
34. **Stormo, G. D.** 1986. Translation intiation, p. 195–224. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, Stoneham, Mass.
35. **Swings, J., and J. DeLey.** 1977. The biology of *Zymomonas*. Bacteriol. Rev. **41**:1–46.
36. **Walker, J. E., A. F. Carne, M. J. Runswick, J. Bridgen, and J. I. Harris.** 1980. Glyceraldehyde-3-phosphate dehydrogenase. Complete amino-acid sequence of the enzyme from *Bacillus stearothermophilus*. Eur. J. Biochem. **108**:549–565.