JOURNAL OF BACTERIOLOGY, Oct. 1992, p. 6548–6553
0021-9193/92/206548-06$02.00/0

Vol. 174, No. 20

Molecular Characterization of the *Zymomonas mobilis* Enolase (*eno*) Gene

MICHEAL E. BURNETT, JIAN LIU, AND TYRRELL CONWAY*

School of Biological Sciences, University of Nebraska, Lincoln, Nebraska 68588-0118

Received 10 June 1992/Accepted 2 August 1992

The *Zymomonas mobilis* gene encoding enolase was cloned by genetic complementation of an *Escherichia coli eno* mutant. An enzyme assay and sodium dodecyl sulfate-polyacrylamide gel electrophoresis confirmed the overexpression of enolase in *E. coli* clones carrying the *Z. mobilis eno* gene. The *eno* gene is present in a single copy on the *Z. mobilis* genome. Nucleotide sequence analysis of the *eno* region revealed an open reading frame of 1,293 bp that encodes a protein of 428 amino acids with a predicted molecular weight of 45,813. Comparison of the sequence of *Z. mobilis* enolase with primary amino acid sequences for other enolases indicates that the enzyme is highly conserved. Unlike all of the previously studied glycolytic genes from *Z. mobilis* that possess canonical ribosome binding sites, the *eno* gene is preceded by a modest Shine-Dalgarno sequence. The transcription initiation site was mapped by primer extension and found to be located within a 115-bp sequence that is 55.7% identical to a highly conserved consensus sequence found within the regulatory regions of highly expressed *Z. mobilis* genes. Northern RNA blot analysis revealed that *eno* is encoded on a 1.45-kb transcript. The half-life of the *eno* mRNA was determined to be 17.7 ± 1.7 min, indicating that it is unusually stable. The abundance of the *eno* message is proposed to account for enolase being the most prevalent protein in *Z. mobilis*.

Zymomonas mobilis is an obligately fermentative, ethanologenic bacterium that uses the Entner-Doudoroff glycolytic pathway exclusively for the catabolism of sucrose, glucose, and fructose, the only carbon sources that support its growth (24, 34). When employed anaerobically, the Entner-Doudoroff pathway is inherently inefficient, yielding only a single mole of ATP per mole of glucose. *Z. mobilis* must convert over 95% of its substrate to carbon dioxide and ethanol to gain sufficient energy for growth (24). This level of metabolic activity is made possible by high-level expression of the glycolytic and ethanologenic enzymes, which account for approximately 50% of the soluble protein (1). Enolase is the most abundant of the glycolytic enzymes in *Z. mobilis*, as determined by two-dimensional gel electrophoresis and biochemical assay (2, 25). It has been suggested that enolase plays a role in metabolic flux control (34), since the intracellular concentrations of the triose phosphates and phosphoglycerates are relatively high (1, 5).

Enolase has been studied extensively in eukaryotic organisms but has received less attention in bacteria (36). The purified enzyme from *Escherichia coli* is dimeric, as in most other organisms, and exhibits kinetic properties similar to those of *Saccharomyces cerevisiae* enolase (32). The purified *Z. mobilis* enolase is unusual in being octomeric but is otherwise similar to the *S. cerevisiae* enzyme (26). Enolase-defective *Pseudomonas aeruginosa* (21) and *E. coli* mutants in which the metabolic lesion leads to accumulation of the phosphoglycerates have been described (16, 18). Part of the nucleotide sequence of the *E. coli eno* gene was determined during a study of the neighboring *pyrG* gene, but this is the only molecular genetic characterization of enolase in a bacterium to date (35). The present study focuses on the cloning and characterization of the *Z. mobilis* gene that encodes enolase. Transcriptional analysis indicates that the *eno* message is remarkably stable, perhaps contributing to the high level of enolase synthesis.

* Corresponding author.

MATERIALS AND METHODS

Bacterial strains and plasmids. *Z. mobilis* CP4 was grown at 30°C in complex medium containing 100 g of glucose per liter as described previously (25). *E. coli* strains were routinely grown at 37°C in Luria broth (23) with no added carbohydrate. *E. coli* DF261 (*tonA22 ompF627 eno-2 relA1 pit-10 spoT1*) (16) was grown at 37°C in M63 minimal medium supplemented to final concentrations of 0.2% glycerol, 0.2% succinate, 0.2% Casamino Acids, and 0.1 mM histidine (33). Complementation of *E. coli* DF261 with a *Z. mobilis* CP4 genomic library (4) was accomplished on minimal medium containing ampicillin (50 mg/ml) and glucose as the sole carbon and energy source (16, 33). One of the enolase-positive clones, pTC144, was selected for study.

DNA manipulations. Transformation, restriction mapping, and subcloning were carried out by standard methods (27, 31). Small-scale plasmid isolations were prepared by the method of Silhavy et al. (31). Restriction enzymes and DNA-modifying enzymes were used according to the recommendations of the manufacturers.

Enzyme assay. *E. coli* and *Z. mobilis* cell extracts were prepared for enzyme assay as described by Conway et al. (8). Enolase activity was assayed spectrophotometrically in a reaction coupled with pyruvate kinase and lactate dehydrogenase by measuring the oxidation of NADH (A_{340}) as described by Pawluk et al. (26).

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Gels were prepared and stained with Coomassie blue as previously described (15).

Southern hybridization analysis. Agarose gels (0.8%) to be used for Southern blotting were electrophoresed and stained with ethidium bromide by standard methods (31). Capillary transfer of DNA to Genescreen Plus hybridization transfer membranes and hybridization were as described previously (10). DNA hybridization probes were prepared by random-primed DNA labeling with a kit.

DNA sequence analysis. The *Z. mobilis eno* gene was sequenced by the dideoxy method (28) with a Sequenase kit and double-stranded DNA sequencing (7). The sequence

was derived from a nested deletion series prepared by digestion with exonuclease III and by using specific oligonucleotide primers. Sequence data were compiled by using the University of Wisconsin Genetic Computer Group sequence analysis software package version 6.1 (11).

RNA isolation and transcriptional analysis. Total RNA was isolated from logarithmic-phase *Z. mobilis* cultures as described previously (15). Northern RNA blot analysis was carried out as described previously (9) with DNA hybridization probes labeled with a random-primed labeling kit. RNA half-life determinations after rifampin addition were conducted as described previously (15). Primer extension analysis of the transcriptional initiation site was accomplished by previously described methods (8). S1 nuclease protection experiments were as described previously (8). The filters were analyzed with the AMBIS Radioanalytic Imaging System (AMBIS Systems, Inc., San Diego, Calif.).

Enzymes and chemicals. Restriction enzymes, DNA-modifying enzymes, and *E. coli* DH5α were obtained from Bethesda Research Laboratories, Inc. (Gaithersburg, Md.). The Sequenase and random-primed DNA labeling kits were ordered from U.S. Biochemical Corp. (Madison, Wis.). An Erase-a-Base kit was purchased from Promega, Inc. (Madison, Wis.). Radioactive compounds were purchased from New England Nuclear Corp. (Boston, Mass.). Biochemicals and coupling enzymes were obtained from Sigma Chemical Co. (St. Louis, Mo.).

RESULTS AND DISCUSSION

Cloning and expression of the *Z. mobilis* enolase gene. The *Z. mobilis* enolase gene was cloned by genetic complementation of *E. coli* DF261. The defective enolase locus renders *E. coli* DF261 incapable of growth on glucose-containing media. A *Z. mobilis* genomic library of 5-kb fragments in pUC18 (4) was transformed into *E. coli* DF261 and plated on minimal glucose medium. Glucose-positive clones were obtained at a frequency of approximately 1 in 250 transformants. Plasmid DNA isolated from 12 clones was found to restore the glucose-positive phenotype in *E. coli* DF261. A biochemical assay was used to confirm that these 12 clones expressed a plasmid-borne enolase gene. Seven of the 12 exhibited greatly (10- to 100-fold) elevated levels of enolase activity relative to the native activity in *E. coli* DH5α. Protein extracts from *E. coli* DH5α(pTC144), the clone exhibiting the highest enolase activity, were run on a denaturing polyacrylamide gel, revealing a prominent band with an apparent mobility of 48,000 Da (Fig. 1). This corresponds precisely to the location of enolase on one-dimensional gels of *Z. mobilis* proteins (2). An additional band of approximately 43,000 Da (identity unknown) was also observed on the gel.

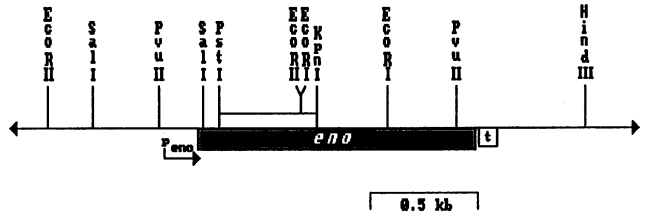

FIG. 2. Restriction map of the *eno* region. The *eno* structural gene is shown below the line. The direction of transcription is indicated by an arrow below the line. The *Pst*I-to-*Kpn*I restriction fragment that was used as a hybridization probe is shown above the line.

A restriction map of the *eno* region was constructed (Fig. 2). Deletions of the *Z. mobilis* DNA insert in pTC144 that eliminated enolase activity were used to localize the *eno* reading frame. A gene-specific *eno* hybridization probe (Fig. 2) was generated and used for Southern analysis of genomic copy number (Fig. 3). The failure of this probe to hybridize to *E. coli* DNA confirmed that the *eno* gene on pTC144 originated from *Z. mobilis* (data not shown). The simple restriction pattern indicated that *eno* is present in a single copy on the *Z. mobilis* genome. This holds true for all of the glycolytic genes studied in *Z. mobilis* thus far (8).

Sequence analysis of the *Z. mobilis eno* region. The nucleotide sequence of a 2,656-bp region of the pTC144 insert was determined, revealing an open reading frame of 1,293 bp capable of encoding a protein of 431 amino acids with a predicted molecular weight of 45,817 (Fig. 4). No additional open reading frames were identified. The deduced primary amino acid sequence of the open reading frame corresponds precisely to the N-terminal sequence of the first 41 amino acids determined for purified enolase (29). The *eno* ribosome binding site consists of the sequence AGG located 7 bp upstream of the ATG initiation codon. This modestly efficient ribosome binding site is unusual for *Z. mobilis* glycolytic genes, which, with the exception of *eno*, all possess canonical ribosome binding sites. In fact, the *eno* ribosome binding site is more similar to those of the *adhB*, *sacA*,

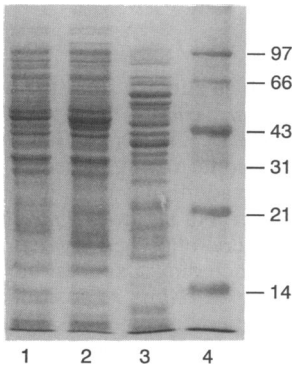

FIG. 1. Cell extracts electrophoresed on a denaturing sodium dodecyl sulfate-polyacrylamide gel stained with Coomassie blue. Lanes: 1, *E. coli* DH5α; 2, *E. coli* DH5α(pTC144); 3, *Z. mobilis* CP4; 4, size markers (kilodaltons). Each lane was loaded with 15 μg of total cell protein.

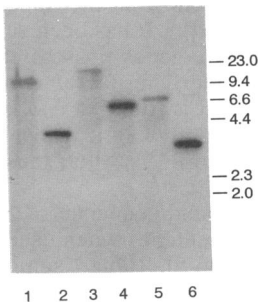

FIG. 3. Southern gel analysis of total *Z. mobilis* genomic DNA with an *eno*-specific DNA hybridization probe. Size markers are in kilobase pairs on the right. DNA was digested with the following (lanes): 1, *Sal*I; 2, *Pst*I; 3, *Kpn*I; 4, *Hin*dIII; 5, *Eco*RV; 6, *Eco*RI.

```
-740  GGC TGG CAA ACG GAT AAA GCG CAT ACG GAA GGC GTG CTG CGA CAC CAA CCA GGC TGG CAA ACG GAT AAA GCG CAT ACG GAA GGC GTG CTG  -651

-650  CGA CAC CAA CCG TCA CAA CTG GGT TAC GGC GGA TAA TGC TAT GAC TGC TTT TGG TAA TGT CAG CCA ATT TTT TAT CTT TCA GAG ATT TTG  -561

-560  ACA ATT TGC CGA TAT GTT TGG AAG CCT TGT GCA TAT AGT CTC CAT AGG AAC TTC CCA GAT TTT CGG CAA CCA ATT CTG GCG GTC TGG TCG  -471

-470  ACG ATT TGG CAC ATT ATT AAC AGC TTC GAT GGC TTT GTG ACG ACC TTC ATC ACC ATT TCA TGA ATT TTT CCA TGA AAG GGA CTG TGA TCC  -381

-380  TGA AAT AAA TGA TGC AGT TTA TTT TCT TCT TTT TTT AGT AAA ATC TTT GGC TTC GGA AAG ATG CGA TTT TGC TTT TTC AGC CAG GTG AGA  -291

-290  AGC CGT TTC CGA TAT ATG CTG TCC TAT ATG ATG GCC TTC ATG TTT TGC TTC ATG AGT GGC TTT TTC AGC CTG ATT TTC GGC GTT TTT CTG  -201

-200  CAA ATC ATT GGC AGC TGA TGC TAC ATT CGC CAC AGT TTT GTT TTC GGC CAT TGT CTA TAC TCC AGT TAC TCA ATA CGT AAC AAT AAT CAG  -111

-110  TTT ATC CTA ACT ATA GAA TCG CAT GAG AAG CGA TAA CGT TTC ACC ATA AGC AAT ATA TTC ATT GCA ACA GTG GAA TTG CCT TAT GCG TCA  -21

 -20  AGG AAG GAT AGA TCA TTG ACG GAC TGA GTT CAA AAA GAG ACT CGT CTA AAA GAT TTT AAG AAA GGT TTC GAT ATG ACA GCT ATT GTC AGT    70
                          +1                                                                  Met Thr Ala Ile Val Ser     6

  71  ATC CAT GGC CGT CAG GTT GTC GAC AGC CGC GGT AAC CCG ACC GTT GAA GTT GAT GTT ACG CTT GAA GAT GGC AGC TTC GGC CGC GCT GCA   160
   7  Ile His Gly Arg Gln Val Val Asp Ser Arg Gly Asn Pro Thr Val Glu Val Asp Val Thr Leu Glu Asp Gly Ser Phe Gly Arg Ala Ala    36

 161  GTG CCG TCA GGT GCT TCT ACC GGC GTT CAT GAA GCT GTT GAA CTT CGT GAT GGC GAC AAA ACC CGT TGG GGT GGT AAA GGC GTC ACC AAA   250
  37  Val Pro Ser Gly Ala Ser Thr Gly Val His Glu Ala Val Glu Leu Arg Asp Gly Asp Lys Thr Arg Trp Gly Gly Lys Gly Val Thr Lys    66

 251  GCT GTT CAC GCT GTA AAC AAC GAA ATT GCT AAC GCA ATT ATT GGT CTG GAA GCC GAA GAT CAG GAA CTG ATC GAC CAG ACG ATG ATC AAG   340
  67  Ala Val His Ala Val Asn Asn Glu Ile Ala Asn Ala Ile Ile Gly Leu Glu Ala Glu Asp Gln Glu Leu Ile Asp Gln Thr Met Ile Lys    96

 341  CTC GAT GGC ACC CCG AAC AAG GGT AAA TTC GGT GCT AAC GCT ATC CTC GGT GTC AGC TTG GCT GTT GCT AAA GCT GCT GCT GAA GCT CGC   430
  97  Leu Asp Gly Thr Pro Asn Lys Gly Lys Phe Gly Ala Asn Ala Ile Leu Gly Val Ser Leu Ala Val Ala Lys Ala Ala Ala Glu Ala Arg   126

 431  GGT CTC CCG CTT TAC CGT TAT GTT GGT GGT ACG GCA GCT CAC GTT CTT CCG GTT CCG ATG ATG AAC ATC GTT AAC GGT GGT ATG CAC GCT   520
 127  Gly Leu Pro Leu Tyr Arg Tyr Val Gly Gly Thr Ala Ala His Val Leu Pro Val Pro Met Met Asn Ile Val Asn Gly Gly Met His Ala   156

 521  GAC AAC CCC ATC GAT TTC CAG GAA TTC ATG ATT GCT CCG GTT GGC GCC AGC TCC ATC AAT GAA GCT GTC CGC ATC GGT ACC GAA GTT TTC   610
 157  Asp Asn Pro Ile Asp Phe Gln Glu Phe Met Ile Ala Pro Val Gly Ala Ser Ser Ile Asn Glu Ala Val Arg Ile Gly Thr Glu Val Phe   186

 611  CAT ACC CTG AAA AAA GAA CTG TCT GCT AAA GGC ATG AAC ACC AAC GTC GGT GAC GAA GGT GGT TTC GCT CCT AGC CTT GAC AGT GCT TCT   700
 187  His Thr Leu Lys Lys Glu Leu Ser Ala Lys Gly Met Asn Thr Asn Val Gly Asp Glu Gly Gly Phe Ala Pro Ser Leu Asp Ser Ala Ser   216

 701  TCT GCT CTG GAC TTC ATC GTC GAT TCC ATC TCC AAA GCC GGT TAC AAG CCG GGC GAA GAT GTG TTC ATC GCT CTC GAT GCA GCT TCC TCC   790
 217  Ser Ala Leu Asp Phe Ile Val Asp Ser Ile Ser Lys Ala Gly Tyr Lys Pro Gly Glu Asp Val Phe Ile Ala Leu Asp Ala Ala Ser Ser   246

 791  GAG TTC TAC AAC AAA GAT CAG AAC ATC TAC GAT CTT AAG GGT GAA GGC CGT AAA CTG ACC TCC GCT CAG CTC GTT GAT TAC TAT GTC GAA   880
 247  Glu Phe Tyr Asn Lys Asp Gln Asn Ile Tyr Asp Leu Lys Gly Glu Gly Arg Lys Leu Thr Ser Ala Gln Leu Val Asp Tyr Tyr Val Glu   276

 881  CTC TGC GGC AAA TAT CCG ATC TAT TCC ATC GAA GAT CTG GCC GAA GAT GAC TTC GAA TTC TGG AAG ATC CTT ACC GAA AAG CTC GGT GAC   970
 277  Leu Cys Gly Lys Tyr Pro Ile Tyr Ser Ile Glu Asp Leu Ala Glu Asp Asp Phe Glu Phe Trp Lys Ile Leu Thr Glu Lys Leu Gly Asp   306

 971  AAA GTT CAG TTG GTC GGT GAC GAT CTG TTC GTG ACC AAC GTG AAG CGT CTT TCT GAT GGT ATC GAA CGC GGT ATC GCC AAC TCG CTG CTC  1060
 307  Lys Val Gln Leu Val Gly Asp Asp Leu Phe Val Thr Asn Val Lys Arg Leu Ser Asp Gly Ile Glu Arg Gly Ile Ala Asn Ser Leu Leu   336

1061  GTG AAG TTC AAC CAG ATC GGT TCT TTG TCT GAA ACG CTC GCA GCC GTT AAC ATG GCT AAC GAC GCT TCT TAC ACG CCT GTT ATG TCT CAC  1150
 337  Val Lys Phe Asn Gln Ile Gly Ser Leu Ser Glu Thr Leu Ala Ala Val Asn Met Ala Asn Asp Ala Ser Tyr Thr Pro Val Met Ser His   366

1151  CGT TCC GTG GAA ACC GAA GAC ACC ACG ATT GCT GAC CTC GCT GTT GCC ACC AAC TGC GGT CAG ATC AAG ACC GGT AGC CTT TGC CGT TCC  1240
 367  Arg Ser Val Glu Thr Glu Asp Thr Thr Ile Ala Asp Leu Ala Val Ala Thr Asn Cys Gly Gln Ile Lys Thr Gly Ser Leu Cys Arg Ser   396

1241  GAA CGT ATC GCT AAA TAC AAT CAG CTG ATG CGC ATC GAA GAA GAA CTG GGT TCG GTT GCT AAA TAT GCT GGC CGT TCG GTT CTT AGA AAA  1330
 397  Glu Arg Ile Ala Lys Tyr Asn Gln Leu Met Arg Ile Glu Glu Glu Leu Gly Ser Val Ala Lys Tyr Ala Gly Arg Ser Val Leu Arg Lys   426

1331  GCC AAA TAA GAA TCA CAG CTA GAA CCG GTT TTT AAA CCG CTC TAA CTG AAT GGG GAC ATT GAG CCT TGA TTT ATC GCT TAA AGG CGG TAA  1420
 427  Ala Lys End                                                                                                          456

1421  TTA TAG GCT CCA TGT CCC TGA TTC GTT CAA AGA ACA GTC TTT TCC GTG CTG CAA TAG GCC CCG CTT GAT GGT TGT TGT CTT GGG GGT CTT  1510

1511  CGT CGG TAA TGC CCT TTT GGG CGA AAT GGT TTA CTG GCA CTC GAT GGT TAC CGC CAA AAA ATG GCC GTT AGA AAA GCG GAA TTG GCG GCT  1600

1601  TTG GAA CAC CAG CGA GAC CAA TTG GCG AAC AGC ATC GCC CTT CTT AAT CCC AAG CCA CTG ATC CTT GAT TAT CCG AAG AAT TGG TGC GTA  1690

1691  AAG AAA CGG GAA ATT CGG CCG GAT GAA GTG ATT GAT CTC GAC AGC CAG AAT AAA TTC AAC TTT TGG AAT GGA TCA CCA TGA TCT AAC GAG  1780

1781  GAC TAA CAG CCT TGT TGA TTA TGG GTA AAA GCT TCT ATT CTA GGC TGA TCG CCG CGA TTG ACA GTT TCT CTT TGA AAG CGT GTC ATT ACT  1870

1871  CAA TAT TCA AGG TCA AAT CTG TCT TGA TGA AGC TTC AGA GTG ACT T                                                          1916
```

FIG. 4. Nucleotide sequence of the *Z. mobilis* enolase gene and flanking regions. The sequence is numbered with respect to the transcriptional initiation site (+1). The ribosome binding site is underlined. The stop codon is indicated (End).

phoC, *trpA*, and *trpB* genes, all of which exhibit low-level expression in *Z. mobilis* (8). It has been suggested that all of the *Z. mobilis* glycolytic genes possess excellent ribosome binding sites as an underlying factor contributing to high-level expression (8). If it is true that *eno* is not translated as efficiently as the other glycolytic genes are, some other factor, such as a high mRNA level, must account for the observed level of enolase.

The deduced amino acid sequence of *Z. mobilis* enolase revealed several close matches when used as a query to the GenBank data base (Fig. 5). Table 1 summarizes the results of the comparisons. Enolase appears to be highly conserved among all organisms studied, with 148 (32.7%) of 453 residues conserved in all species. The *Z. mobilis* enzyme most closely resembles the *E. coli* enolase, with 95 of the first 132 residues conserved. Interestingly, the *Z. mobilis* enzyme more closely resembles enolase from the higher eukaryotes than those from *S. cerevisiae*. The *Z. mobilis* enolase is the most divergent of the sequences studied. Nevertheless, the three ligands of the conformational cation within the active site region and four additional residues in the active site region are all conserved in the *Z. mobilis* enzyme (20). Several deletions were observed by comparison with *S. cerevisiae* enolase, including a six-residue deletion at position 295; two-residue deletions at positions 2, 88, and 214; and one-residue deletions at positions 311 and 372 (Fig. 5). These deletions were fitted to the three-dimensional structure of *S. cerevisiae* enolase derived from the coordinates of Lebioda et al. (20) by molecular modeling with an Iris 4D/320DGX system; no conformational perturbations of the active-site region were predicted by the computer.

Transcriptional analysis of *eno*. The 5′ end of the *eno*

```
            1                                                                                                  100
Zymomonas   M..TAIVSIH GRQVVDSRGN PTVEVDVTLE DGSFGRAAVP SGASTGVHEA VELRDGDKTR WGGKGVTKAV HAVNNEIANA IIGLEAE..D QELIDQTMIK
E. coli     M..SKIVKII GREIIDSRGN PTVEAEVHLE GGFVGMAAAP SGASTGSREA LELRDGDKSR FLGKGVTKAV AAVNGPIAQA LIGKDAK..D QAGIDKIMID
Yeast       MAVSK.VY.A .RSVYDSRGN PTVEVELTTE KGVF.RSIVP SGASTGVHEA LEMRDEDKSK WMGKGVMNAV NNVNNVIAAA FVKANLDVKD QKAVDDFLLS
Maize       MAVTITWVKA .RQIFDSRGN PTVEVDVGLS DGSYARGAVP SGASTGIYEA LELRDGGSDY L.GKGVLKAV SNVNNIIGPA IVGKDPTEQV EIDNFMVQQL
Insect      M..TIKAIKA .RQIYDSRGN PTVEVDLTTE LGLF.RAAVP SGASTGVHEA LELRDNDKAN YHGKSVLKAV GHVNDTLGPE LIKANLDVVD QASIDNFMIK
Amphibian   M..SIKNIRA .REIFDSRGN PTVEVDLYTC KGLF.RAAVP SGASTGIYEA LELRDNDKTR YLGKGVGRAV KYVNEFLGPA LCTQNLNVVE QEKIDKLMIE
Avian       M..SILKIHA .REIFDSRGN PTVEVDLYTN KGLF.RAAVP SGASTGIYEA LELRDNDKTR YMGKGVSKAV EHINKTIAPA LISKNVNVVE QDKIDKLMLD
Human       M..SILKIHA .REIFDSRGN PTVEVDLFTS KGLF.RAAVP SGASTGIYEA LELRDNDKTR YMGKGVSKAV EHINKTIAPA LVSKKLNVTE QEKIDKLMIE
  Consensus M---------- -R---DSRGN PTVE------ -G-------P SGASTG--EA -E-RD----- --GK-V--AV ---N------ ---------- ----------

            101                                                                                                200
Zymomonas   LDGTPNKG.. ..KFGANAIL GVSLAVAKAA AEARGLPLYR YVGGTAAH.. ...VLPVPMM NIVNGGMHAD NPIDFQEFMI APVGASSINE AVRIGTEVFH
E. coli     LDGTEKKS.. ..KFGANAIL AVSLANAKAA A......... .......... .......... .......... .......... .......... ..........
Yeast       LDGTANKS.. ..KLGANAIL GVSMAAARAA AAEKNVPLYQ HLADLSKSKT SPYVLPVPFL NVLNGGSHAG GALAIQEFMI APTGAKTFAE AMRIGSEVYH
Maize       DGTSNEWGWC KQKLGANAIL AVSLAVCKAG AMVKKIPLYQ HIANLAGNKT L..VLPVPAF NVINGGSHAG NKLAMQEFMI LPTGASSFKE AMKMGVEVYH
Insect      LDGTENKS.. ..KFGANAIL GVSLAVAKAG AAKKGVPLYK HIADLAGNKE I..ILPVPAF NVINGGSHAG NKLAMQEFMI LPTGATSFTE AMKMGSEVYH
Amphibian   MDGTENKS.. ..KFGANALL GVSLAVCKAG AAEKGVPLYR HIADLAGNPE V..ILPVPAF NVINGGSHAG NKLAMQEFMI LPVGADSFKE AMRIGAEVYH
Avian       MDGSENKS.. ..KFGANAIL GVSLAVCKAG AAEKGVPLYR HIADLAGNPE V..ILPVPAF NVINGGSHAG NKLAMQEFMI PPCGADSFKE AMRIGAEVYH
Human       MDGTENKS.. ..KFGANAIL GVSLAVCKAG AVEKGVPLYR HIADLAGNSE V..ILPVPAF NVINGGSHAG NKLAMQEFMI LPVGAANFRE AMRIGAEVYH
  Consensus ---------- --K-GANA-L -VS-A---A- A-----PLY- ---------- ----LPVP-- N--NGG-HA- -----QEFMI -P-GA----E A---G-EV-H

            201                                                                                                300
Zymomonas   TLKKELSAKG MNT..NVGDE GGFAPSLDSA SSALDFIVDS ISKAGYKPGE DVFIALDAAS SEFYNKDQNI YDLKGEGRKL TSAQLVDYYV ELCG.....K
E. coli     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Yeast       NLKSLTKKRY GASAGNVGDE GGVAPNIQTA EEALDLIVDA IKAAGHDGKV KI..GLDCAS SEFFKDGKYD LD..FKNPES DKSKWLTGVE LADMYHSLMK
Maize       NLKSIIKKKY GQDATNVGDE GGFAPNIQEN KEGLELLKAA IEKAGYTGKV VI..GMDVAA SEFFGEKDKT YDLNFKEENN DGSNKISGDS LKDLYKSFVS
Insect      HLKNVIKAKF GLDATAVGDE GGFAPNIQSN KEALNLISDA IAKAGYTGKI EI..GMDVAA SEFYKDGQYD LD..FKNEKS DKSQWLPADK LANLYKEFIK
Amphibian   NLKNVIKEKY GKDATNVGDE GGFAPNILEN KEALELLKTA INKAGYPDKI VI..GMDVAA SEFYRDGKYD LD..FKSPDD PSRYISP.DK LAELYMSFVK
Avian       NLKNVIKEKY GKDATNVGDE GGFAPNILEN KEALELLKTA IGKAGYSDKV VI..GMDVAA SEFYRDGKYD LD..FKSPDD PSRYISP.DQ LADLYKGFVK
Human       NLKNVIKEKY GKDATNVGDE GGFAPNILEN KEGLELLKTA IGKAGYTDKV VI..GMDVAA SEFFRSGKYD LD..FKSPDD PSRYISP.DQ LADLYKSFIK
  Consensus -LK------- -----NVGDE GG-AP----- ---L------ I--AG----- ------D-A- SEF------- -D-------- ---------- ----------

            301                                                                                                400
Zymomonas   .YPIYSIEDL .AEDDFEFWK ILTEKLGDKV QLVGDDLFVT NVKRLSDGIE RGIANSLLVK FNQIGSLSET L.AAVNMAND ASYTPVMSHR SVETEDTTIA
E. coli     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Yeast       RYPIVSIEDP FAEDDWEAWS HFFKTAGI.. QIVADDLTVT NPARIATAIE KKAADALLLK VNQIGTLSES IKAAQDSFAA NWGVMV.SHR SGETEDTFIA
Maize       EYPIESIEDP FDQDDWSTYA KLTDEIGQKV QIVGDDLLVT NPTRVAKAIN EKTCNALLLK VNQIGSVTES IEAVRMSKRA GWGVMA.SHR SGETEDTFIA
Insect      DFPIVSIEDP FDQDHWEAWS NLTGCTDI.. QIVGDDLTVT NPKRIATAVE KKACNCLLLK VNQIGTVTES IAAHLLAKKN GWGTMV.SHR SGETEDSFIG
Amphibian   NYPVVSIEDP FDQDHWEAWT KFTAASGI.. QVVGDDLTVT NPKRIAKAVE EKACNCLLLK VNQIGTVTES LEACKLAQSN GWGVMV.SHR SGETEDTFIA
Avian       NYPVVSIEDP FDQDDWGAWK KFTGSVGI.. QVVGDDLTVT NPKRIAKAVE EKACNCLLLK VNQIGSVTES LQACKLAQSN GWGVMV.SHR SGETEDTFIA
Human       DYPVVSIEDP FDQDDWGAWQ KFTASAGI.. QVVGDDLTVT NPKRIAKAVN EKSCNCLLLK VNQIGSVTES LQACKLAQAN GWGVMV.SHR SGETEDTFIA
  Consensus --P--SIED- ---D------ ---------- Q-V-DDL-VT N--R------ ------LL-K -NQIG---E- ---------- -------SHR S-ETEDT-I-

            401                                                453
Zymomonas   DLAVATNCGQ IKTGSLCRSE RIAKYNQLMR IEEELGSVAK YAGRSVLRKA K*.
E. coli     .......... .......... .......... .......... .......... ...
Yeast       DLVVGLRTGQ IKTGAPARSE RLAKLNQLLR IEEELGDKAV YAGENFHHGD KL*
Maize       DLSVGLSTGQ IKTGAPCRSE RLAKYNQLLR IEEELGDAAV YAGAKFRAPV EPY
Insect      DLVVGLSTGQ IKTGAPCRSE RLAKYNQILR IEEEIGAGVK FAGKSFGKPQ *..
Amphibian   DLVVGLCTGQ IKTGAPCRSE RLAKYNQLLR IEEELGSKAR FAGKNFRKPV FN*
Avian       DLVVGLCTGQ IKTGAPCRSE RLAKYNQLLR IEEELGSKAR FAGRNFRNPR IN*
Human       DLVVGLCTGQ IKTGAPCRSE RLAKYNQLLR IEEELGSKAK FAGRNFRNPL AK*
  Consensus DL-V----GQ IKTG---RSE R-AK-NQL-R IEEE-G---- -AG------- ---
```

FIG. 5. Comparison of the *Z. mobilis* enolase amino acid sequence with those of *E. coli* (35), *S. cerevisiae* (17), maize (*Zea mays* [19]), an insect (*Drosophila melanogaster* [6]), an amphibian (*Xenopus laevis* [30]), an avian (*Anas platyrhynchos*, Peking duck [36]), and human neuron-specific enolase-1 (14). Dots denote gaps introduced to maximize identities. Numbering is with respect to the consensus line.

mRNA was determined by primer extension analysis (Fig. 6). This result was confirmed by S1 nuclease protection analysis (data not shown). Transcriptional initiation proceeds from a guanine residue located 53 bp upstream of the *eno* start codon. A comparison of the sequence immediately upstream of the mapped transcriptional initiation site with the previously described consensus derived for highly expressed *Z. mobilis* glycolytic genes (3) revealed a significant

TABLE 1. Enolase amino acid comparisons

Species	% Identity of enolases						
	Z. mobilis	*E. coli*[a]	*S. cerevisiae*	Maize	Insect	Amphibian	Avian
Z. mobilis							
E. coli	72.5						
S. cerevisiae	49.9	53.8					
Maize	48.6	45.5	56.5				
Insect	52.8	57.6	63.8	61.6			
Amphibian	53.4	59.8	61.6	64.2	67.5		
Avian	54.3	60.6	62.5	64.7	77.0	91.3	
Human	54.1	68.7	69.5	69.5	76.4	90.0	93.4

[a] Only the first 132 amino acids were available for this protein.

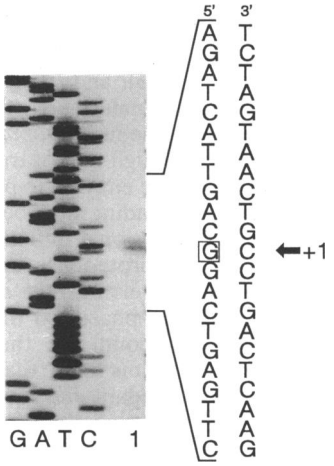

FIG. 6. Localization of the *eno* promoter by primer extension analysis. Sequencing lanes are labeled G, A, T, and C. The sequence of the region is shown on the right for clarity (arrow). The +1 denotes the start of the *eno* promoter sequence. Lane 1 carries the primer extension product.

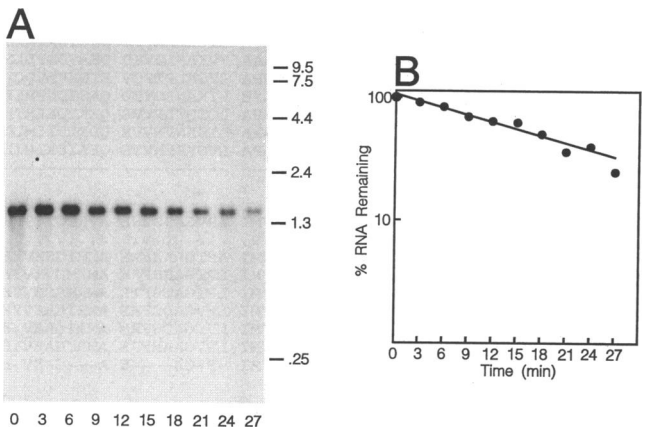

FIG. 7. (A) Transcript stability determined by Northern blot analysis of the *Z. mobilis eno* transcript with a gene-specific hybridization probe (Fig. 2). Total RNA was harvested from log-phase cultures after inhibition of transcription at the times indicated (in minutes) below the gel. Gels were loaded with 2.5 μg of RNA per lane. Size markers (in kilobases) are provided at the right. (B) The data shown in panel A were quantitated by radioanalytic scanning of the filters and are shown on a semilogarithmic plot. The decay rate was determined by measuring the slope of the line. The half-life of the *eno* mRNA was 17.7 ± 1.7 min.

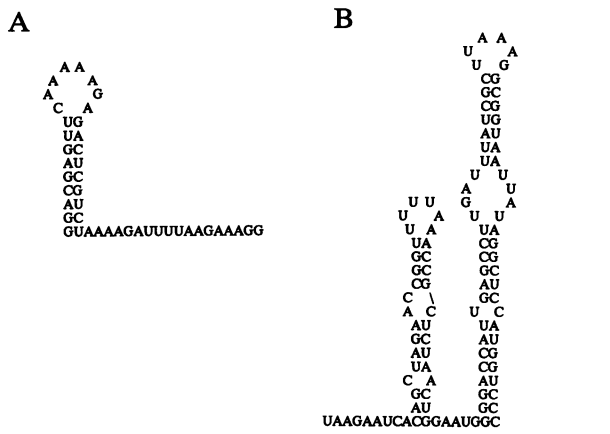

FIG. 8. (A) The 5′ end of the *eno* message is predicted to form a stem-loop in which the ultimate 5′ end is base paired within the stem. (B) The 3′ end of the *eno* message is predicted to form these two adjacent stem-loop structures.

match (55.7% identity). The strong identity of the *eno* promoter region with the *Z. mobilis* consensus implies efficient transcription initiation. It should be noted that enolase has been shown to be synthesized at a rate greater than that of the bulk of cell protein (25).

Northern blot analysis showed that *eno* is transcribed on a 1.45-kb transcript, indicating that *eno* is monocistronic (Fig. 7A). The *eno* message appears to be highly abundant, requiring just 30 min of exposure to X-ray film for 2.5 μg of total RNA. Also, the *eno* message is present well into the stationary phase (data not shown). The prevalence of the *eno* transcript could be the result of rapid transcription and/or an unusually slow degradation rate. To consider this latter possibility, the decay rate of the *eno* message was determined (Fig. 7B). The measured decay rate is 17.7 ± 1.7 min, which indicates that *eno* is more stable than other *Z. mobilis* transcripts, including two transcripts of the *gap-pgk* operon (9), *gap* at 16 min (12), *gap-pgk* at 7 min (12), *pgi* at 6.0 min (15), *glf-zwf-edd-glk* at 12 min (3), and *frk* at 6.2 min (37).

Factors contributing to high enolase activity. The relative abundance of the enolase enzyme in *Z. mobilis* might be due to high levels of *eno* mRNA, a high rate of translation of the *eno* message, or resistance of the enzyme to proteolysis. The structure of the *eno* ribosome binding site does not appear to be conducive to efficient translational initiation. Resistance of the glycolytic enzymes to proteolysis appears to be a general property, since high levels of most of the enzymes remain well into the stationary phase, so this factor alone would not be expected to account for the high enolase activity (2, 25). Therefore, it seems likely that the high level of enolase results from an abundant *eno* message, perhaps due to a combination of rapid transcription and significant mRNA stability. Stability of the *eno* message might be imparted by the substantial stem-loop structures located both upstream and downstream of the *eno* structural gene (Fig. 8), as has been suggested for other *Z. mobilis* genes (22). Computer analysis (University of Wisconsin Genetics Computer Group TERMINATOR program) did not predict a strong transcriptional terminator downstream of the gene. However, two prominent stem-loop structures were predicted in the region 100 bases downstream of the stop codon (Fig. 8B). The first stem-loop structure, spanning bases 1346 to 1378, has an upstream stem of 14 bases, a loop of 6 bases, and a downstream stem of 13 bases with a total of 11 paired bases (University of Wisconsin Genetics Computer Group FOLD program). The second stem-loop structure, spanning bases 1383 to 1438, has a stem of 26 bases with 21 bases paired and a loop of 6 bases (FOLD program). The measured length of the 1.45-kb mRNA is in keeping with the length that would be predicted based on the mapped location of the transcriptional initiation site (Fig. 6) and possible termination immediately downstream of the larger of the two stem-loop structures. It is also possible that termination actually occurs at a site further downstream and that the major stem-loop structure serves to stabilize the 3′ end of the *eno* message by blocking further 3′-to-5′ exonuclease degradation, as has been suggested for other *Z. mobilis* mRNAs (22). Moreover, the 5′ end of the *eno* message (Fig. 8A) is predicted by computer (FOLD program) to be paired in a simple stem-loop structure, a feature of the 5′ region of *E. coli* mRNAs that has been shown to impart significant stability (13). Interestingly, other stable *Z. mobilis* mRNAs, including the processed monocistronic *glf* and *glk* transcripts, also possess upstream stem-loop structures in which the ultimate 5′ ends of the messages are paired (22). The relative contributions of transcription rate and degradation rate to high *eno* mRNA levels, and hence high enolase activity, are being investigated.

ACKNOWLEDGMENTS

Computer time was provided courtesy of the University of Nebraska, Lincoln, Computer Resources Center. This work was supported by grants from the U.S. Department of Energy (DE-FG02-90ER20002) and the Nebraska Energy Office (UNR-11).

We thank B. Bachmann for providing strains and J. Stezowski and P. Goebel for assisting with three-dimensional enzyme structure analysis.

REFERENCES

1. **Algar, E. M., and R. Scopes.** 1985. Studies on cell-free metabolism: ethanol production by extracts of *Zymomonas mobilis*. J. Biotechnol. **2**:275–287.

2. **An, H., R. K. Scopes, M. Rodriguez, K. F. Mackenzie, 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. **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.
4. **Barnell, W. O., K. C. Yi, and T. Conway.** 1990. Sequence and genetic organization of a *Zymomonas mobilis* gene cluster that encodes several enzymes of glucose metabolism. J. Bacteriol. **172:**7227–7240.
5. **Barrow, K. D., J. G. Collins, R. S. Norton, P. L. Rogers, and G. M. Smith.** 1984. ^{31}P nuclear magnetic resonance studies of the fermentation of glucose to ethanol by *Z. mobilis*. J. Biol. Chem. **259:**5711–5716.
6. **Bishop, J. G., and V. G. Corces.** 1990. The nucleotide sequence of a *Drosophila melanogaster* enolase gene. Nucleic Acids Res. **18:**191–199.
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., R. Fliege, D. Jones-Kilpatrick, J. Liu, W. O. Barnell, and S. E. Egan.** 1991. Cloning and characterization of the *Zymomonas mobilis eda* gene that encodes 2-keto-3-deoxy-6-phosphogluconate aldolase of the Entner-Doudoroff pathway. Mol. Microbiol. **5:**2901–2911.
9. **Conway, T., and L. O. Ingram.** 1989. Phosphoglycerate kinase gene for *Zymomonas mobilis*: cloning, sequencing, and localization within the *gap* operon. J. Bacteriol. **170:**1926–1933.
10. **Conway, T., and L. O. Ingram.** 1989. Similarity of *Escherichia coli* propanediol oxidoreductase (*fucO* product) and an unusual alcohol dehydrogenase from *Zymomonas mobilis* and *Saccharomyces cerevisiae*. J. Bacteriol. **171:**3754–3759.
11. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis tools for the VAX. Nucleic Acids Res. **12:**287–395.
12. **Eddy, C. K., J. P. Mejia, T. Conway, and L. O. Ingram.** 1989. Differential expression of *gap* and *pgk* genes within the *gap* operon of *Zymomonas mobilis*. J. Bacteriol. **173:**6549–6554.
13. **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.
14. **Giallongo, A., S. Feo, R. Moore, C. M. Croce, and L. C. Showe.** 1986. Molecular cloning and nucleotide sequence of a full length cDNA for human alpha enolase. Proc. Natl. Acad. Sci. USA **83:**6741–6745.
15. **Hesman, T. L., W. O. Barnell, and T. Conway.** 1991. Cloning, characterization, and nucleotide sequence analysis of a *Zymomonas mobilis* phosphoglucose isomerase gene that is subject to carbon source-dependent regulation. J. Bacteriol. **173:** 3215–3223.
16. **Hillman, J. D., and D. G. Fraenkel.** 1975. Glyceraldehyde 3-phosphate dehydrogenase mutants of *Escherichia coli*. J. Bacteriol. **122:**1175–1179.
17. **Holland, M. J., J. P. Holland, G. P. Thill, and K. A. Jackson.** 1981. The primary structure of two yeast enolase genes: homology between the 5' noncoding flanking regions of yeast enolase and glyceraldehyde-3-phosphate dehydrogenase genes. J. Biol. Chem. **256:**1385–1395.
18. **Irani, M. H., and P. K. Maitra.** 1977. Properties of *Escherichia coli* mutants deficient in enzymes of glycolysis. J. Bacteriol. **132:**398–410.
19. **Lal, S. K., S. Johnson, T. Conway, and P. M. Kelly.** 1991. Characterization of a maize cDNA that complements an enolase-deficient mutant of *Escherichia coli*. Plant Mol. Biol. **16:** 787–795.
20. **Lebioda, L., B. Stec, and J. M. Brewer.** 1989. The structure of yeast enolase at 2.25-Angstrom resolution. J. Biol. Chem. **264:**3685–3693.
21. **Lessie, T. G., and P. V. Phibbs.** 1984. Alternative pathways of carbohydrate utilization in pseudomonads. Annu. Rev. Microbiol. **38:**359–387.
22. **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.
23. **Luria, S. E., and M. Delbruck.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28:**491–511.
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 Inc., Menlo Park, Calif.
25. **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.
26. **Pawluk, A., K. Scopes, and K. Griffiths-Smith.** 1986. Isolation and properties of the glycolytic enzymes from *Zymomonas mobilis*. J. Biochem. **238:**275–281.
27. **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.
28. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
29. **Scopes, R. K. (La Trobe University, Bundoora, Australia).** Personal communication.
30. **Segil, N., A. Shrutkowski, M. B. Dworkin, and E. Dworkin-Rastl.** 1988. Enolase isoenzymes in adult and developing *Xenopus laevis* and characterization of a cloned enolase sequence. Biochem. J. **251:**31–39.
31. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Spring, T. G., and F. Wold.** 1971. The purification and characterization of *Escherichia coli* enolase. J. Biol. Chem. **246:**6797–6802.
33. **Thomson, J., P. D. Gerstenberger, D. E. Goldberg, E. Gociar, A. O. deSilva, and D. G. Fraenkel.** 1979. ColE1 hybrid plasmids for *Escherichia coli* genes of glycolysis and the hexose monophosphate shunt. J. Bacteriol. **137:**502–506.
34. **Viikari, L.** 1988. Carbohydrate metabolism in *Zymomonas mobilis*. Crit. Rev. Biotechnol. **7:**237–261.
35. **Weng, M., C. A. Makaroff, and H. Zalkin.** 1986. Nucleotide sequence from *E. coli pyrG* gene encoding CTP synthetase. J. Biol. Chem. **261:**5568–5574.
36. **Wold, F.** 1971. Enolase, p. 499–538. *In* P. D. Boyer (ed.), The enzymes, vol. 5, 3rd ed. Academic Press, Inc., New York.
37. **Zembruski, B., P. Chilco, S. L. Liu, J. Liu, T. Conway, and R. K. Scopes.** 1991. The fructokinase gene from *Zymomonas mobilis*: cloning, sequencing, expression and standard comparison with other hexose kinases. J. Bacteriol. **174:**3455–3460.