JOURNAL OF BACTERIOLOGY, July 1992, p. 4638–4646
0021-9193/92/144638-09$02.00/0
Copyright © 1992, American Society for Microbiology
Vol. 174, No. 14

Molecular Characterization of the Entner-Doudoroff Pathway in *Escherichia coli*: Sequence Analysis and Localization of Promoters for the *edd-eda* Operon

SARA E. EGAN,[1] RONDA FLIEGE,[1] SUXIANG TONG,[1] ANNEMARIE SHIBATA,[1] RICHARD E. WOLF, JR.,[2] AND TYRRELL CONWAY[1]*

School of Biological Sciences, University of Nebraska, Lincoln, Nebraska, 68588-0118,[1] and Department of Biological Sciences, University of Maryland Baltimore County, Catonsville, Maryland 21288[2]

Received 21 February 1992/Accepted 18 May 1992

The nucleotide sequence of the entire *Escherichia coli edd-eda* region that encodes the enzymes of the Entner-Doudoroff pathway was determined. The *edd* structural gene begins 236 bases downstream of *zwf*. The *eda* structural gene begins 34 bases downstream of *edd*. The *edd* reading frame is 1,809 bases long and encodes the 602-amino-acid, 64,446-Da protein 6-phosphogluconate dehydratase. The deduced primary amino acid sequences of the *E. coli* and *Zymomonas mobilis* dehydratase enzymes are highly conserved. The *eda* reading frame is 642 bases long and encodes the 213-amino-acid, 22,283-Da protein 2-keto-3-deoxy-6-phosphogluconate aldolase. This enzyme had been previously purified and sequenced by others on the basis of its related enzyme activity, 2-keto-4-hydroxyglutarate aldolase. The data presented here provide proof that the two enzymes are identical. The primary amino acid sequences of the *E. coli*, *Z. mobilis*, and *Pseudomonas putida* aldolase enzymes are highly conserved. When *E. coli* is grown on gluconate, the *edd* and *eda* genes are cotranscribed. Four putative promoters within the *edd-eda* region were identified by transcript mapping and computer analysis. P_1, located upstream of *edd*, appears to be the primary gluconate-responsive promoter of the *edd-eda* operon, responsible for induction of the Entner-Doudoroff pathway, as mediated by the *gntR* product. High basal expression of *eda* is explained by constitutive transcription from P_2, P_3, and/or P_4 but not P_1.

The Entner-Doudoroff pathway consists of two enzymes, 6-phosphogluconate dehydratase, encoded by the *edd* gene, and 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase, encoded by the *eda* gene (11). In *Escherichia coli*, this pathway is employed for gluconate metabolism (10, 15). Dehydratase activity is virtually absent in cells grown on glucose and is strictly inducible for growth on gluconate (10, 41). On the other hand, high basal levels of KDPG aldolase activity are found regardless of the carbon source (12, 28). Although aldolase synthesis is constitutive, threefold induction occurs for growth on gluconate and twofold induction occurs for growth on hexuronic acids. The latter is consistent with the role of KDPG aldolase as a key enzyme of the peripheral pathway for hexuronic acid metabolism, a function independent of the Entner-Doudoroff pathway (28).

The molecular biology of the Entner-Doudoroff pathway has been investigated in some detail in *Zymomonas mobilis* (2, 5), but few molecular studies of the Entner-Doudoroff pathway in *E. coli* have been conducted (14). In *E. coli*, the *edd* and *eda* genes are tightly linked to the *zwf* gene, which codes for glucose-6-phosphate dehydrogenase (13). The location of these genes on the *E. coli* physical map was reported recently (7). Despite tight linkage, the three related genes are controlled differently. The complete nucleotide sequence of the *E. coli zwf* gene has been published, and the promoter has been located (30). Genetic and physical analyses of *zwf* expression have established that the *zwf* gene is monocistronic and subject to growth rate-dependent regulation (29). Expression of *edd*, as well as gluconate transport and gluconokinase, is controlled in a negative fashion by the product of *gntR*, which presumably encodes a repressor protein (19). Expression of *eda* is subject to control by the *kdgR* product, as are the permease and kinase for 2-keto-3-deoxygluconate (28). In this report, we demonstrate that induction of the Entner-Doudoroff pathway at the biochemical level is paralleled by derepression of the *edd* and *eda* genes and that these two genes of the Entner-Doudoroff pathway can be cotranscribed. Furthermore, we report on the molecular structure of the *edd-eda* region and the locations of four transcriptional start sites for *edd* and *eda*.

* Corresponding author.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The strains and plasmids used in this study are listed in Table 1. *E. coli* strains were routinely grown in Luria broth (24), with or without added carbohydrate (0.5%), at 37°C. All experimental cultures were harvested in the mid-logarithmic phase. Antibiotic-resistant transformants were selected by addition of ampicillin (50 mg/liter). Media were made solid by addition of agar (1.5%). Constructions containing both orientations of the intact *edd* gene, designated pTC180 and pTC181, were prepared by subcloning a 3.1-kb *Sal*I-to-*Pst*I fragment from pDR2, in accordance with the published restriction map (30), into pBluescriptII (KS+ and SK+). Selection for the plasmid-borne, functional *edd* gene was accomplished by complementation of defective gluconate metabolism in *E. coli* RW231 (38) on minimal medium as previously described (2, 35, 36). Plasmid pLC37-44 was transferred into *E. coli* RW231 via conjugal mating as already described (30). Constructions containing the intact *eda* gene in opposite orientations, pTC190 and pTC191, were made by subcloning of a 1.7-kb *Nru*I restriction fragment (see Fig. 7) from pLC37-44 (7) into the *Eco*RV site of pBluescriptII. Selection of the functional *eda* gene was

TABLE 1. Plasmids and strains used

Strain or plasmid	Relevant genotype	Source or reference
E. coli strains		
DH5αF′	F^{+} *lacZ*M15 *recA*	BRL[a]
DF214	Δ(*eda-zwf*)*15 pgi-7*::Mu	36
RW231	Δ(*eda-zwf*)*209* Δ(*sbcB-rfb*)*210 kdgR50*	38
Plasmids		
pUC18	*bla lacI′Z′*[b]	40
pBluescriptII	*bla lacI′Z′* f1 origin	Stratagene
pDR2	*zwf edd*	30
pLC37-44	*edd eda*	4
pTC180	*edd*	This study
pTC181	*edd*	This study
pTC190	*eda*	This study
pTC191	*eda*	This study
pTC196	*eda′*	This study

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

accomplished by complementation of a defect in glucuronic acid metabolism in *E. coli* DF214 as described previously (5). A subclone of a portion of the *eda* region, pTC196, was made by ligating a 0.3-kb *Hin*cII-to-*Pst*I restriction fragment (see Fig. 7) into pBluescriptII.

Enzyme assays. *E. coli* cells were prepared for enzyme assays as described previously (2). 6-Phosphogluconate dehydratase was assayed by published methods (41). KDPG aldolase was assayed as described previously (5).

DNA methods. Transformation and recombinant DNA techniques were carried out by using standard methods (31). Small-scale plasmid isolations were done by a modification of the method of Birnboim and Doly (3) as previously described (31). Restriction enzymes and DNA-modifying enzymes were used as recommended by the manufacturers.

DNA sequence analysis. The *E. coli edd* and *eda* genes were sequenced by the dideoxy method (32) by using Sequenase. Sequencing strategies were based on creation of nested deletion series of pTC180, pTC181, pTC190, and pTC191, which were amplified for single-stranded templates in *E. coli* DH5αF′ as described previously (2). Sequence data were analyzed by using the University of Wisconsin Genetics Computer Group sequence analysis software package, version 6.1 (9).

RNA isolation and analysis. RNA isolations were conducted as described previously (18). Northern (RNA) analysis was carried out as previously described (6), by using DNA hybridization probes that were labeled with a randomly primed labeling kit. An *edd*-specific DNA hybridization probe was prepared as a 1.14-kb *Bst*EII restriction fragment from pTC180. An *eda*-specific probe was prepared from pTC196 as a 0.28-kb *Hin*cII-to-*Pst*I restriction fragment. Determinations of RNA sizes were based on migration of known RNA standards, as described previously (18). Primer extension analysis of transcriptional initiation sites was accomplished as described previously (5). The oligonucleotide used for analysis of *edd* was 26 bases long, spanning nucleotides 1718 to 1743 (Fig. 1). The oligonucleotide used for analysis of *eda* was 22 bases long, spanning nucleotides 3578 to 3599 (Fig. 1). S1 nuclease protection experiments were used to confirm primer extension results, as described previously (5).

Enzymes and chemicals. Restriction enzymes, murine leukemia virus reverse transcriptase, and DNA modifying enzymes were obtained from Bethesda Research Laboratories, Inc. (Gaithersburg, Md.). Sequenase was ordered from U.S. Biochemical Corp. (Cleveland, Ohio). Radioactive compounds were purchased from DuPont NEN Products (Boston, Mass.). Biochemicals were obtained from Sigma Chemical Co. (St. Louis, Mo.).

Nucleotide sequence accession numbers. The entire nucleotide sequence of the *edd-eda* region is available in the GenBank data base, under accession no. M87458. The nucleotide sequence of the *zwf* gene has already been published (30) and is available in the GenBank data base, under accession no. M55005.

RESULTS

Sequence analysis of the *edd-eda* region. Plasmid clones containing the *E. coli edd* and *eda* genes were prepared on the basis of the known locations of these genes on published restriction maps (7, 30). Genetic complementation of appropriate *E. coli* mutants confirmed that the genes were intact and functional. This finding was supported by biochemical assay of the two Entner-Doudoroff enzymes in crude extracts of *E. coli* RW231(pTC180) and DF214(pTC190). The nucleotide sequence of the *E. coli edd-eda* region is shown in Fig. 1. The reported sequence begins with a *Stu*I site that lies 79 bp downstream of the *zwf* stop codon. It was necessary to include a portion of the published *zwf* sequence here, beginning at the *Stu*I site, because the *edd-eda* regulatory sequences lie in this region. The first base in the *Stu*I palindrome is numbered 1556, in accordance with the published *zwf* sequence (30). The *edd-eda* sequence continues to the *Bam*HI site downstream of *eda*, at bp 4490.

(i) *edd* gene. The *E. coli edd* gene is located 236 bp downstream of *zwf*. The *edd* structural gene is 1,809 bp long, corresponding to a reading frame of 603 codons. The ATG start codon is preceded (5 bp upstream) by a modest ribosome-binding site, AGAG, that is predicted to provide only moderate translational initiation efficiency (33). The reading frame ends with the stop codon, TAA. A computer search of the sequence downstream of the *edd* region failed to reveal any transcriptional terminatorlike structures (39; University of Wisconsin Genetics Computer Group TERMINATOR program). The deduced amino acid sequence corresponds to a protein of 602 amino acids with a molecular weight of 64,446. An alignment of *E. coli* 6-phosphogluconate dehydratase with the *Z. mobilis* enzyme (2) is shown in Fig. 2. The two dehydratase enzymes are quite similar, exhibiting 57.7% identity, and 75.3% of the residues are chemically conserved. These enzymes are both related by primary sequence to the *E. coli* dihydroxy-acid dehydratase that is encoded by the *ilvD* gene (8). An alignment of the two *E. coli* enzymes shows them to be 30.3% identical and 54.2% conserved (Fig. 2). A comparison of *Z. mobilis* 6-phosphogluconate dehydratase and *E. coli* dihydroxy-acid dehydratase has already been reported (2).

(ii) *eda* gene. The *E. coli eda* gene begins just 34 bp downstream of *edd*. The ATG start codon is preceded (9 bp) by the ribosome-binding site, GAGAG. As was the case for the *edd* gene, maximally efficient translational initiation from this ribosome-binding site would not be expected (33). The *eda* reading frame is 642 bp long and ends with the stop codon, TAA. A strong transcriptional terminatorlike structure, beginning 10 bp downstream of the structural gene, was identified by computer search (University of Wisconsin Genetics Computer Group TERMINATOR program). It is

```
                                                                          P1>
1556 AGGC CTC GGT GCC GAT TCA CCC ACG AGG CTT TTT TTA TTA CAC TGA CTG AAA CGT TTT TGC CCT ATG AGC TCC GGT TAC AGG CGT TTC AGT   1645

1646 CAT AAA TCC TCT GAA TGA AAC GCG TTG TGA ATC ATC CTG CTC TGA CAA CTC AAT TTC AGA GCC TTT ATG AAT CCA CAA TTG TTA CGC GTA   1735
   1                                                                               edd M   N   P   Q   L   L   R   V      8

1736 ACA AAT CGA ATC ATT GAA CGT TCG CGC GAG ACT CGC TCT GCT TAT CTC GCC CGG ATA GAA CAA GCG AAA ACT TCG ACC GTT CAT CGT TCG   1825
   9 T   N   R   I   I   E   R   S   R   E   T   R   S   A   Y   L   A   R   I   E   Q   A   K   T   S   T   V   H   R   S     38

1826 CAG TTG GCA TGC GGT AAC CTG GCA CAC GGT TTC GCT GCC TGC CAG CCA GAA GAC AAA GCC TCT TTG AAA AGC ATG TTG CGT AAC AAT ATC   1915
  39 Q   L   A   C   G   N   L   A   H   G   F   A   A   C   Q   P   E   D   K   A   S   L   K   S   M   L   R   N   N   I     68

1916 GCC ATC ATC ACC TCC TAT AAC GAC ATG CTC TCC GCG CAC CAG CCT TAT GAA CAC TAT CCA GAA ATC ATT CGT AAA GCC CTG CAT GAA GCG   2005
  69 A   I   I   T   S   Y   N   D   M   L   S   A   H   Q   P   Y   E   H   Y   P   E   I   I   R   K   A   L   H   E   A     98

2006 AAT GCG GTT GGT CAG GTT GCG GGC GGT GTT CCG GCG ATG TGT GAT GGT GTC ACC CAG GGG CAG GAT GGA ATG GAA TTG TCG CTG CTA AGC   2095
  99 N   A   V   G   Q   V   A   G   G   V   P   A   M   C   D   G   V   T   Q   G   Q   D   G   M   E   L   S   L   L   S    128

2096 CGC GAA GTG ATA GCG ATG TCT GCG GCG GTG GGG CTG TCC CAT AAC ATG TTT GAT GGT GCT CTG TTC CTC GGT GTG TGC GAC AAG ATT GTC   2185
 129 R   E   V   I   A   M   S   A   A   V   G   L   S   H   N   M   F   D   G   A   L   F   L   G   V   C   D   K   I   V    158

2186 CCG GGT CTG ACG ATG GCA GCC CTG TCG TTT GGT CAT TTG CCT GCG GTG TTT GTG CCG TCT GGA CCG ATG GCA AGC GGT TTG CCA AAT AAA   2275
 159 P   G   L   T   M   A   A   L   S   F   G   H   L   P   A   V   F   V   P   S   G   P   M   A   S   G   L   P   N   K    188

2276 GAA AAA GTG CGT ATT CGC CAG CTT TAT GCC GAA GGT AAA GTG GAC CGC ATG GCC TTA CTG GAG TCA GAA GCC GCG TCT TAC CAT GCG CCG   2365
 189 E   K   V   R   I   R   Q   L   Y   A   E   G   K   V   D   R   M   A   L   L   E   S   E   A   A   S   Y   H   A   P    218

2366 GGA ACA TGT ACT TTC TAC GGT ACT GCC AAC ACC AAC CAG ATG GTG GTG GAG TTT ATG GGG ATG CAG TTG CCA GGC TCT TCT TTT GTT CAT   2455
 219 G   T   C   T   F   Y   G   T   A   N   T   N   Q   M   V   V   E   F   M   G   M   Q   L   P   G   S   S   F   V   H    248

2456 CCG GAT TCT GCC TGC GAT GCT TTG ACC GCC GCA GCT GCG CGT CAG GTT ACA CGC ATG ACC GGT AAT GGT AAT GAA TGG ATG CCG ATC GGT   2545
 249 P   D   S   A   C   D   A   L   T   A   A   A   A   R   Q   V   T   R   M   T   G   N   G   N   E   W   M   P   I   G    278

2546 AAG ATG ATC GAT GAG AAA GTG GTG GTG AAC GGT ATC GTT GCA CTG CTG GCG ACC GGT GGT TCC ACT AAC CAC ACC ATG CAC CTG GTG GCG   2635
 279 K   M   I   D   E   K   V   V   V   N   G   I   V   A   L   L   A   T   G   G   S   T   N   H   T   M   H   L   V   A    308

2636 ATG GCG CGC GCG GCC GGT ATT CAG ATT AAC TGG GAT GAC TTC TCT GAC CTT TCT GAT GTT GTA CCG CTG ATG GCA CGT CTC TAC CCG AAC   2725
 309 M   A   R   A   A   G   I   Q   I   N   W   D   D   F   S   D   L   S   D   V   V   P   L   M   A   R   L   Y   P   N    338

2726 GGT CCG GCC GAT ATT AAC CAC TTC CAG GCG GCA GGT GGC GTA CCG GTT CTG GTG CGT GAA CTG CTC AAA GCA GGC CTG CTG CAT GAA GAT   2815
 339 G   P   A   D   I   N   H   F   Q   A   A   G   G   V   P   V   L   V   R   E   L   L   K   A   G   L   L   H   E   D    368

2816 GTC AAT ACG GTG GCA GGT TTT GGT CTG TCT CGT TAT ACC CTT GAA CCA TGG CTG AAT AAT GGT GAA CTG GAC TGG CGG GAA GGG GCG GAA   2905
 369 V   N   T   V   A   G   F   G   L   S   R   Y   T   L   E   P   W   L   N   N   G   E   L   D   W   R   E   G   A   E    398

2906 AAA TCA CTC GAC AGC AAT GTG ATC GCT TCC TTC GAA CAA CCT TTC TCT CAT CAT GGT GGG ACA AAA GTG TTA AGC GGT AAC CTG GGC CGT   2995
 399 K   S   L   D   S   N   V   I   A   S   F   E   Q   P   F   S   H   H   G   G   T   K   V   L   S   G   N   L   G   R    428

2996 GCG GTT ATG AAA ACC TCT GCC GTG CCG GTT GAG AAC CAG GTG ATT GAA GCG CCA GCG GTT GTT TTT GAA AGC CAG CAT GAC GTT ATG CCG   3085
 429 A   V   M   K   T   S   A   V   P   V   E   N   Q   V   I   E   A   P   A   V   V   F   E   S   Q   H   D   V   M   P    458

3086 GCC TTT GAA GCG GGT TTG CTG GAC CGC GAT TGT GTC GTT GTT GTC CGT CAT CAG GGG CCA AAA GCG AAC GGA ATG CCA GAA TTA CAT AAA   3175
 459 A   F   E   A   G   L   L   D   R   D   C   V   V   V   V   R   H   Q   G   P   K   A   N   G   M   P   E   L   H   K    488
                                         P2>
3176 CTC ATG CCG CCA CTT GGT GTA TTA TTG GAC CGG TGT TTC AAA ATT GCG TTA GTT ACC GAT GGA CGA CTC TCC GGC GCT TCA GGT AAA GTG   3265
 489 L   M   P   P   L   G   V   L   L   D   R   C   F   K   I   A   L   V   T   D   G   R   L   S   G   A   S   G   K   V    518

3266 CCG TCA GCT ATC CAC GTA ACA CCA GAA GCC TAC GAT GGC GGG CTG CTG GCA AAA GTG CGC GAC GGG GAC ATC ATT CGT GTG AAT GGA CAG   3355
 519 P   S   A   I   H   V   T   P   E   A   Y   D   G   G   L   L   A   K   V   R   D   G   D   I   I   R   V   N   G   Q    548

3356 ACA GGC GAA CTG ACG CTG CTG GTA GAC GAA GCG GAA CTG GCT GCT CGC GAA CCG CAC ATT CCT GAC CTG AGC GCG TCA CGC GTG GGA ACA   3445
 549 T   G   E   L   T   L   L   V   D   E   A   E   L   A   A   R   E   P   H   I   P   D   L   S   A   S   R   V   G   T    578
                                                                                             P3>         P4>
3446 GGA CGT GAA TTA TTC AGC GCC TTG CGT GAA AAA CTG TCC GGT GCC GAA CAG GGC GCA ACC TGT ATC ACT TTT TAA GAC GAC AAA TTT GTA   3535
 579 G   R   E   L   F   S   A   L   R   E   K   L   S   G   A   E   Q   G   A   T   C   I   T   F   END                      602

3536 ATC AGG CGA GAG AAA ACT CTG ATG AAA AAC TGG AAA ACA AGT GCA GAA TCA ATC CTG ACC ACC GGC CCG GTT GTA CCG GTT ATC GTG GTA   3625
   1                         eda M   K   N   W   K   T   S   A   E   S   I   L   T   T   G   P   V   V   P   V   I   V   V     23

3626 AAA AAA CTG GAA CAC GCG GTG CCG ATG GCA AAA GCG TTG GTT GCT GGT GGG GTG CGC GTT CTG GAA GTG ACT CTG CGT ACC GAG TGT GCA   3715
  24 K   K   L   E   H   A   V   P   M   A   K   A   L   V   A   G   G   V   R   V   L   E   V   T   L   R   T   E   C   A     53

3716 GTT GAC GCT ATC CGT GCT ATC GCC AAA GAA GTG CCT GAA GCG ATT GTG GGT GCC GGT ACG GTG CTG AAT CCA CAG CAG CTG GCA GAA GTC   3805
  54 V   D   A   I   R   A   I   A   K   E   V   P   E   A   I   V   G   A   G   T   V   L   N   P   Q   Q   L   A   E   V     83

3806 ACT GAA GCG GGT GCA CAG TTC GCA ATT AGC CCG GGT CTG ACC GAG CCG CTG CTG AAA GCT GCT ACC GAA GGG ACT ATT CCT CTG ATT CCG   3895
  84 T   E   A   G   A   Q   F   A   I   S   P   G   L   T   E   P   L   L   K   A   A   T   E   G   T   I   P   L   I   P    113

3896 GGG ATC AGC ACT GTT TCC GAA CTG ATG CTG GGT ATG GAC TAC GGT TTG AAA GAG TTC AAA TTC TTC CCG GCT GAA GCT AAC GGC GGC GTG   3985
 114 G   I   S   T   V   S   E   L   M   L   G   M   D   Y   G   L   K   E   F   K   F   F   P   A   E   A   N   G   G   V    143

3986 AAA GCC CTG CAG GCG ATC GCG GGT CCG TTC TCC CAG GTC CGT TTC TGC CCG ACG GGT GGT ATT TCT CCG GCT AAC TAC CGT GAC TAC CTG   4075
 144 K   A   L   Q   A   I   A   G   P   F   S   Q   V   R   F   C   P   T   G   G   I   S   P   A   N   Y   R   D   Y   L    173

4076 GCG CTG AAA AGC GTG CTG TGC ATC GGT GGT TCC TGG CTG GTT CCG GCA GAT GCG CTG GAA GCG GGC GAT TAC GAC CGC ATT ACT AAG CTG   4165
 174 A   L   K   S   V   L   C   I   G   G   S   W   L   V   P   A   D   A   L   E   A   G   D   Y   D   R   I   T   K   L    203

4166 GCG CGT GAA GCT GTA GAA GGC GCT AAG CTG TAA AAG TCA AAA TTC CCG ATC GAG GAT CGG GAA TTT TTG TAG CGT TTT TTT GCC GGG TTT   4255
 204 A   R   E   A   V   E   G   A   K   L   END                                                                             213

4256 AAC CCT GTA CTT TTA CCT GTC CGG CGG CGT CGT TCG CGC GTT CAA TGG CGT CAA CAA CAC TCT CTG CAG TAG CCA GTG CCA CCC CCA GAC   4345

4346 GAC GGC TGC CAT CAA TTT CCG GCT TAC CAA ATA AAC GAA TCT GCA AAT CTG CGC TAC GGC ATT CTG CAC ATT ATC AAA CGT GAC ATT CTG   4435

4436 ACT GGT CAG TTG TGG CAG AAT AAC GGC AGA AGC TGC AGG ACC ATA CTG ACG GAT CC                                                4490
```

FIG. 1. Nucleotide sequence of the *E. coli edd-eda* region. The sequence begins at the *Stu*I site located between *zwf* and *edd*. The numbering begins with base 1556 of the published *zwf* sequence (30). The beginnings of both reading frames are labeled with the gene designations. The ribosome-binding sites are underlined. The stop codons are labeled "END." The reported sequence ends at the *Bam*HI site. The −10 and −35 regions of the four putative promoters are underlined, and the transcriptional initiation sites are indicated over the initiation bases (>). The potential transcriptional terminator downstream of *eda* is overlined.

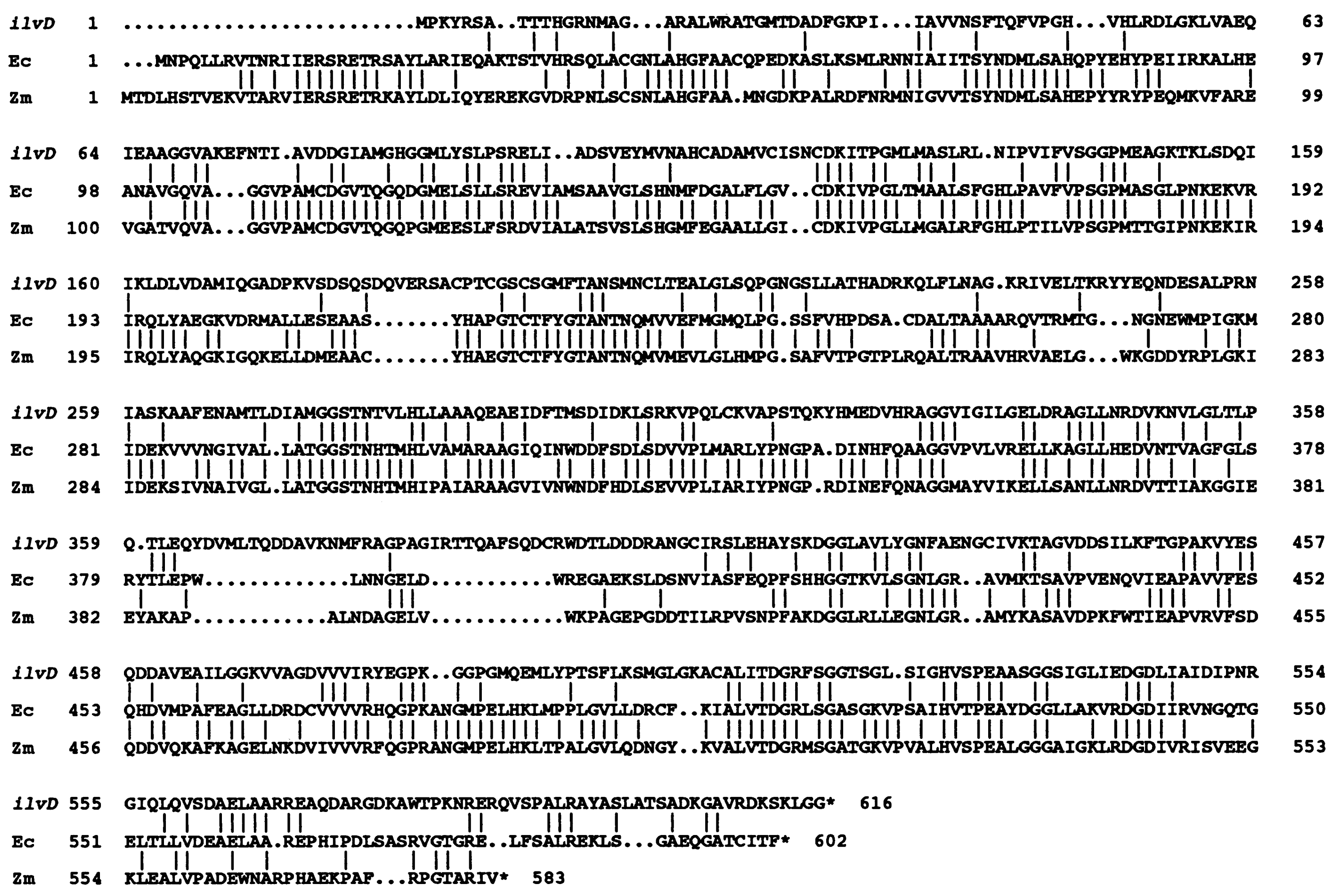

FIG. 2. Amino acid comparisons of the deduced *E. coli edd* gene product (Ec) with the *Z. mobilis edd* product (Zm) and the *E. coli ilvD* product (ilvD). Identity is indicated by vertical lines, and gaps are indicated by periods.

likely that all transcripts covering the *edd-eda* region stop at this site, corresponding to nucleotide 4232, since no other potential terminators were identified. The deduced amino acid sequence of KDPG aldolase corresponds to a protein of 213 amino acids with an aggregate molecular weight of 22,283. A comparison of this peptide sequence to the Swiss Protein data base revealed perfect identity to *E. coli* 2-keto-4-hydroxyglutarate (KHG) aldolase (37). The complete amino acid sequence of KHG aldolase indicated that the N-terminal methionine is retained. In fact, the gene that encodes KHG aldolase has been cloned from *E. coli* by using the polymerase chain reaction and degenerate oligonucleotides, on the basis of the published amino acid sequence (27). The nucleotide sequence reported by Patil and Dekker (27) is indeed identical to the *eda* sequence shown in Fig. 1, indicating that KDPG aldolase is a multifunctional enzyme. KDPG aldolase is able to catalyze interconversion of KDPG with glyceraldehyde-3-phosphate and pyruvate, interconversion of KHG with pyruvate and glyoxylate, and β-decarboxylation of oxaloacetate (27). An alignment of *E. coli* KDPG aldolase with the same enzyme from *Z. mobilis* (5) and *Pseudomonas putida* (34) is shown in Fig. 3. The *E. coli* enzyme is 51.5% identical and 68.4% similar to the *Z. mobilis* enzyme and 44.4% identical and 65.4% similar to the *P. putida* enzyme. Active-site lysine 133 (34) and active-site arginine 49 (37) are conserved in all three aldolases.

Transcriptional analysis of *edd* and *eda* in *E. coli* W3110. Transcriptional regulation of *edd* and *eda* in wild-type *E. coli* W3110 was investigated by Northern hybridization analysis (Fig. 4). Cultures were grown in Luria broth containing 0.5% carbohydrate and harvested in the mid-logarithmic phase, and total RNA was extracted. Growth on gluconate, but not growth on glucose or glucuronic acid, induced transcription of a 2.6-kb message that could be detected with either the *edd*- or *eda*-specific hybridization probe. This transcript was of sufficient length to encode both genes (7). The *eda*-specific hybridization probe revealed two additional transcripts of 1.0 and 0.75 kb that were present in cells grown on each of the three different carbon sources. The ratio of the 0.75-kb transcript to the 1.0-kb transcript increased with growth on gluconate by comparison with growth on glucose or glucuronic acid.

Mapping of the 5′ ends of the *edd-eda* transcripts. The 5′ ends of mRNAs covering the *E. coli edd-eda* region were mapped by primer extension analysis and nuclease S1 protection. A single 5′ transcript end was identified by primer extension of an oligonucleotide spanning the upstream end of the *edd* structural gene (Fig. 5A). The cDNA extension product was 140 bases long, placing the transcriptional initiation site for *edd* 109 bp upstream of the *edd* start codon, bp 1605 in Fig. 1. This result was confirmed by nuclease S1 analysis by using an end-labeled 0.87-kb *Sal*I-to-*Sph*I restriction fragment as a hybridization probe (see Fig. 7). A 229-bp protected fragment was detected following S1 nuclease digestion of the DNA-RNA hybrid (Fig. 5B). This corresponds precisely to the 5′ transcript end revealed by primer

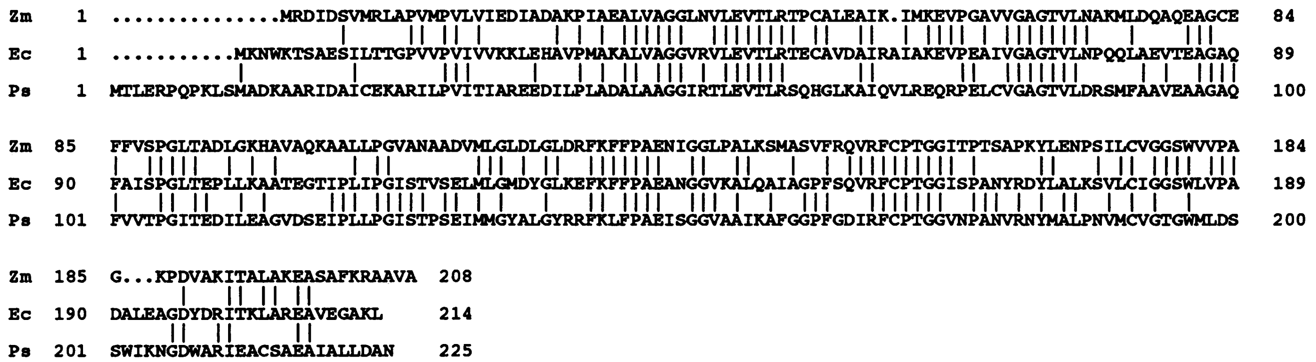

FIG. 3. Amino acid comparisons of the deduced *E. coli eda* gene product (Ec) with the *Z. mobilis eda* product (Zm) and the *P. putida eda* product (Ps). Identity is indicated by vertical lines, and gaps are indicated by periods.

extension and is labeled P_1 in Fig. 1 and 4. A minor 5′ end, mapping to bp 1610 (Fig. 1), was also identified by both methods. No other 5′ transcript ends were observed in these experiments.

Transcript mapping of the *eda* gene revealed several 5′ ends clustered in two regions, both within the upstream *edd* structural gene. Primer extension of an oligonucleotide spanning the upstream end of the *eda* gene resulted in major cDNAs that were 79 and 86 bases long, as well as several much longer extension products, the longest of which was 396 bases long (Fig. 6A). The locations of the two *eda*-proximal 5′ transcript ends were bp 3515, within the downstream end of the *edd* structural gene, and bp 3530, within the *edd-eda* intergenic region (Fig. 1). The longest of the extension products indicated a transcriptional start site at bp 3205, well within the *edd* structural gene. Additional cDNAs, with lengths of approximately 290 bases, were observed (Fig. 6A), but these were not confirmed by S1 nuclease protection. An end-labeled 1.30-kb *Pvu*II restriction fragment covering a large segment of the *edd* structural gene, the intergenic region, and a portion of the upstream

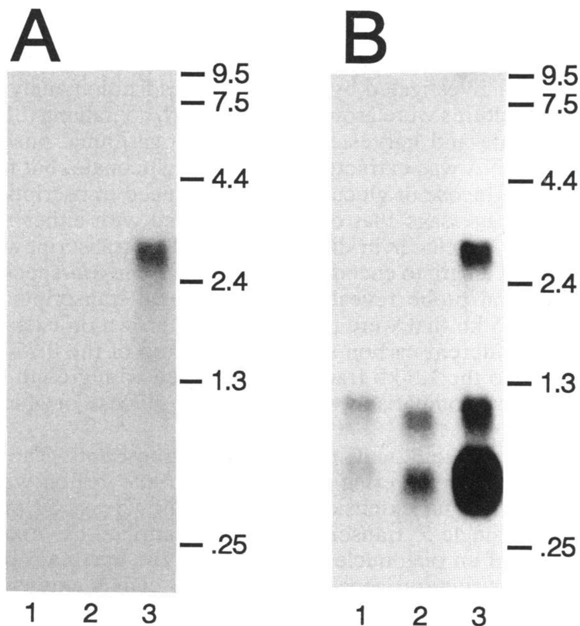

FIG. 4. Northern blot analysis of *edd-eda* expression in *E. coli* W3110. Total RNA was harvested from log-phase cultures grown in complex medium containing glucose (lanes 1), glucuronic acid (lanes 2), or gluconate (lanes 3). The gels were loaded with 5.0 μg of RNA per lane. RNA size standards (in kilobases) are shown on the right of both panels. (A) Hybridization with an *edd*-specific probe. (B) Hybridization with an *eda*-specific probe.

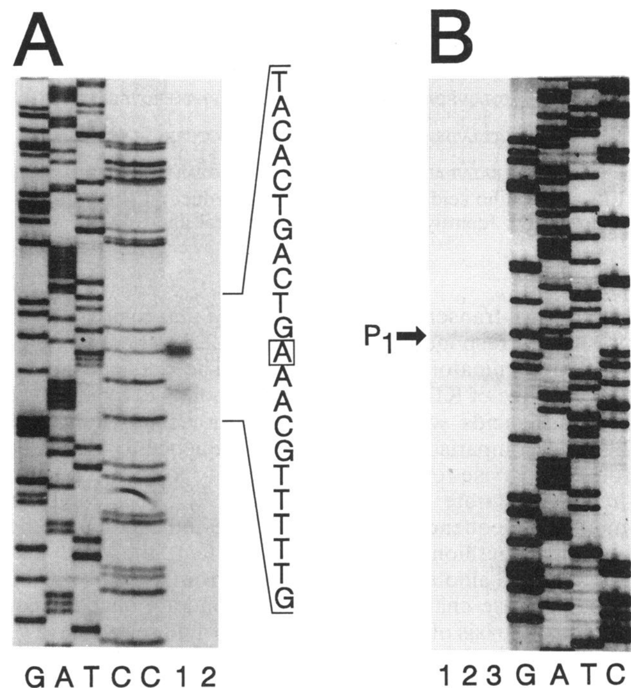

FIG. 5. 5′-end mapping of the *edd-eda* operon mRNA (P_1). (A) Primer extension analysis using an oligonucleotide covering the upstream region of *edd*. The sequence ladder (labeled G, A, T, and C) was generated by using the same oligonucleotide as that used for primer extension. The primer extension reaction was run in lane 1. The sequence of the transcriptional initiation site is detailed on the right (as the complementary sequence, for easy comparison to Fig. 1), and the 5′ end is boxed. A control using yeast tRNA is shown in lane 2. (B) S1 nuclease protection to confirm the primer extension result. Details of the experiment are provided in Results. The same sequence ladder as that used for panel A is labeled. Each lane contained 20 μg of *E. coli* total RNA. The samples were treated with 50 U (lane 3), 100 U (lane 2), or 200 U (lane 1) of S1 nuclease. The protected fragment (P_1) is shown.

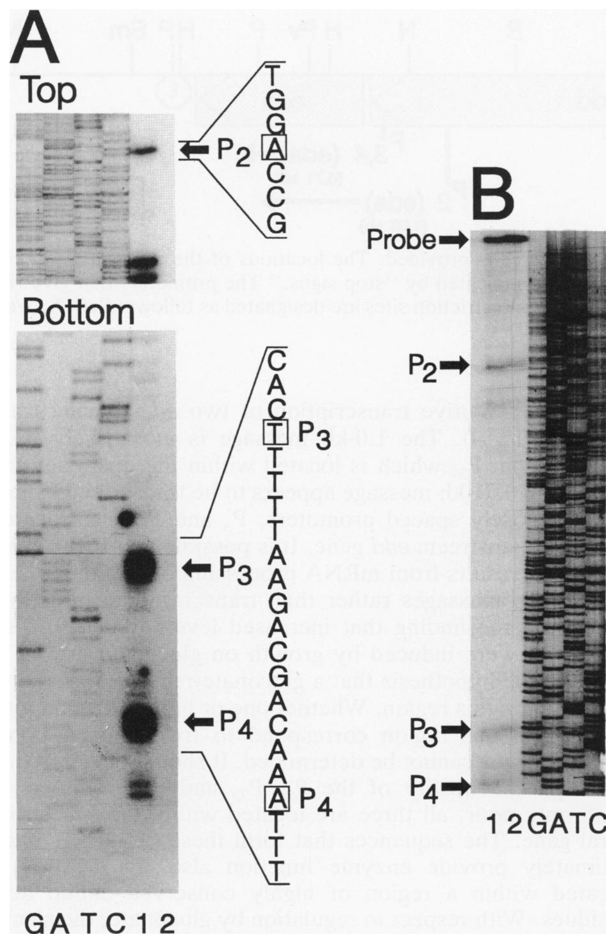

FIG. 6. 5′-end mapping of the *eda*-specific mRNAs. (A) Primer extension analysis using an oligonucleotide covering the upstream region of *eda*. The sequence ladder (labeled G, A, T, and C) was generated by using the same oligonucleotide as that used for primer extension. Only the top and bottom portions of the (same) gel are shown. The primer extension reaction was run in lane 1. The sequence of the transcriptional initiation sites is detailed on the right (as the complementary sequence, for easy comparison to Fig. 1), and the 5′ ends are boxed. The extension products corresponding to P_2, P_3, and P_4 are indicated. A control using yeast tRNA is shown in lane 2. (B) S1 nuclease protection to confirm the primer extension results. Details of the experiment are provided in Results. The same sequence ladder as that used for panel A is labeled. Each lane contained 20 μg of *E. coli* total RNA. The samples were treated with 50 U (lane 2) or 100 U (lane 1) of S1 nuclease. The protected fragments corresponding to P_2, P_3, and P_4 are indicated.

end of *eda* was used as a DNA hybridization probe in an S1 nuclease protection experiment (Fig. 6B). The two *eda*-proximal transcriptional start sites were confirmed by S1 nuclease protection analysis, which resulted in protected DNA-RNA hybrids of 280 and 265 bases, corresponding precisely to the 5′ ends mapped by primer extension (Fig. 6B). An additional protected DNA-RNA hybrid of approximately 585 bases was detected, confirming the longest of the primer extension products shown in Fig. 6A. Thus, 5′ transcript ends in the *eda* region were identified in two locations. One transcriptional start site, labeled P_2 in Fig. 1 and 5, was located 354 bp upstream of the *eda* start codon, well within the upstream *edd* gene. The two additional transcriptional start sites are labeled P_3 and P_4 in Fig. 1 and 5, one within the far downstream end of *edd* and the other immediately downstream of the *edd* structural gene, located 44 and 29 bp, respectively, upstream of the *eda* start codon.

After mapping of the locations of 5′ transcript ends, it was possible to predict the lengths of mRNAs covering the *edd-eda* region based on the basis of transcriptional termination at the terminatorlike structure described above. This information is presented in the model shown in Fig. 7. Transcription from P_1 would give rise to a 2.63-kb mRNA, corresponding to a transcript of this length that was identified by Northern hybridization to both *edd*- and *eda*-specific DNA probes (Fig. 4). Transcription from P_2 would give rise to a 1.03-kb transcript, in keeping with the size of the *eda*-specific transcript observed on Northern blots (Fig. 4). Transcription from P_3 and P_4 would give rise to two transcripts of approximately 0.71 kb that would be indistinguishable on Northern blots. A transcript of this approximate length was also revealed by Northern analysis with the *eda*-specific hybridization probe (Fig. 4).

Computer searches for potential *edd-eda* promoter sequences. The primer extension and S1 nuclease protection experiments described above do not allow distinction between 5′ transcript ends that represent true sites of transcriptional initiation from ends generated by mRNA processing events. To identify sequences that could serve as promoter elements responsible for transcriptional initiation from the mapped transcript ends, the nucleotide sequence shown in Fig. 1 was subjected to computer analysis. Computer programs were used to conduct searches based on the total information content of the sequences and a back-propagation neural network (26) trained to identify promoters of the 16-, 17-, and 18-bp spacing classes (25). The results of these searches are presented in Table 2. A promoter sequence of the 17-bp spacing class that could transcribe *edd* was identified immediately upstream of the 5′ end in this region, spanning nucleotides 1571 to 1599. This promoter is labeled P_1 in Fig. 1 and 5. The −10 and −35 regions of P_1 show a 50% match to the consensus (25). A second promoter sequence of the 18-bp spacing class, spanning nucleotides 3172 to 3201, that could give rise to the transcript end in this region was identified and is labeled P_2 in Fig. 1 and 6. The −10 region of P_2 is a good match to the consensus, while the −35 region is a poor match (25). Two potential promoter sequences that could give rise to the transcript ends in the downstream region of the *edd* gene were identified by computer searching. Both of these promoters are of the 16-bp spacing class. The promoter labeled P_3 in Fig. 1 and 6 spans nucleotides 3480 to 3507, and P_4 spans nucleotides 3490 to 3517. Neither of these promoter sequences is a good match to the consensus sequence (25). Since the levels of the 2.6- and 0.75-kb transcripts are apparently regulated by growth on gluconate, it seemed reasonable to predict that the regulatory sequences in the P_1 and the P_3-to-P_4 regions would contain a common sequence element that would be involved in regulation by gluconate. A computer comparison of the regions revealed a 9-bp sequence, CGGTGCCGA, in the vicinity of the −35 regions of these promoters that was not found elsewhere within the reported sequence. Neither of these sequence elements is in a region of dyad symmetry, as would be expected of control sites (23). Definitive proof that any of these putative promoter sequences are involved in transcription of the *edd-eda* region requires further characterization.

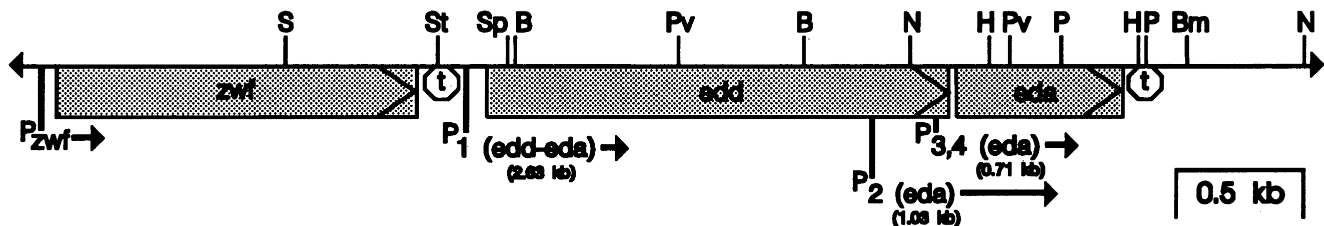

FIG. 7. Restriction map of the *E. coli zwf edd eda* region. A 0.5-kb size marker is provided. The locations of the *zwf*, *edd*, and *eda* structural genes are indicated. Terminators (t) identified by computer searches are indicated by "stop signs." The promoters that give rise to the *zwf* transcript (30) and the *edd-eda* transcripts (Fig. 4) are indicated by arrows. Restriction sites are designated as follows: *Sal*I, S; *Stu*I, St; *Sph*I, Sp; *Bst*EII, B; *Pvu*II, Pv; *Nru*I, N; *Hin*cII, H; *Pst*I, P; *Bam*HI, Bm.

DISCUSSION

Sequence analysis of the *E. coli edd-eda* region allowed identification of both reading frames. Functional activity of the genes was confirmed biochemically by enzyme assay and genetically by complementation of *E. coli* mutants with specific defects in carbon metabolism. With the addition of this new information, the nucleotide sequence of the entire *zwf edd eda* region has been determined. The *edd* structural gene lies 236 bp downstream of *zwf*, and the *eda* structural gene lies 34 bp downstream of *edd*. The *zwf* region has been subjected to extensive molecular characterization (29, 30). The *zwf* gene is monocistronic, and a search of the *zwf-edd* intergenic region for terminators revealed a likely structure, located immediately upstream of P_1 for *edd-eda* (30). Transcriptional analysis provided strong evidence for cotranscription of *edd* and *eda* when cultures are grown on gluconate. A search of the entire *edd-eda* nucleotide sequence for transcriptional terminatorlike structures revealed only one, found immediately downstream of *eda*. This fact, together with the short intergenic region, provides still further evidence that *edd* and *eda* form an operon that functions for induction of the Entner-Doudoroff pathway when *E. coli* is grown on gluconate.

The current results are entirely consistent with those of numerous previous studies that have provided genetic evidence for differential, yet interestingly symmetrical, expression of the *edd* and *eda* genes (14). Molecular characterization of *edd-eda* transcription revealed the presence of four putative promoters in this region. A model describing transcription of the *zwf*, *edd*, and *eda* genes, based on the results of this and related studies (29, 30), is shown in Fig. 7. The 2.6-kb message transcribed from P_1 encodes both *edd* and *eda*. The fact that this transcript was observed only in cells grown on gluconate argues strongly that P_1 is regulated by gluconate. Thus, it is hypothesized that P_1 is the primary gluconate-inducible promoter responsible for induction of the Entner-Doudoroff pathway, as mediated by the *gntR* product (19).

The high basal levels of *eda* expression were shown to be due to constitutive transcription of two *eda*-specific transcripts (Fig. 4). The 1.0-kb message is most likely transcribed from P_2, which is located within the upstream *edd* gene. The 0.71-kb message appears to be transcribed from a pair of closely spaced promoters, P_3 and P_4, also located within the upstream *edd* gene. It is possible that the 0.71-kb transcript results from mRNA processing of the longer 2.6- and 1.0-kb messages rather than transcriptional initiation. However, the finding that increased levels of the 0.71-kb transcript were induced by growth on gluconate argues in favor of the hypothesis that a gluconate-regulated promoter is present in this region. Whether one or both of the mapped 5′ ends in this region correspond to true transcriptional initiation sites cannot be determined. It should be noted that although the quality of the P_2, P_3, and P_4 promoters is uniformly poor, all three are located within the *edd* structural gene. The sequences that form these promoters must ultimately provide enzyme function also. In fact, P_2 is located within a region of highly conserved amino acid residues. With respect to regulation by gluconate, the significance of a conserved nine-base sequence overlapping the −35 regions of gluconate-responsive promoters P_1, P_3, and P_4 is not known (Fig. 1). There is a need for further characterization of the putative *edd-eda* promoter regions before the molecular details of induction of the Entner-Doudoroff pathway can be fully understood.

The *E. coli edd* gene is only the second to be sequenced, in addition to the *Z. mobilis edd* gene (2). A comparison of the deduced primary amino acid sequences of 6-phosphogluconate dehydratase from the two organisms showed that they are highly conserved (Fig. 2). It is interesting that 6-phosphogluconate dehydratase is related to dihydroxyamino acid dehydratase, an enzyme of the isoleucine-valine biosynthetic pathway, which catalyzes a similar reaction. Although 6-phosphogluconate dehydratase has been purified and its kinetics have been examined (22), little is known about the structure of the protein as it relates to reaction mechanism (16). In a recent study, it was shown that *E. coli* 6-phosphogluconate dehydratase is sensitive to superoxide and that the substrate, 6-phosphogluconate, protects the enzyme from attack by superoxide (16). Considering the protective role of 6-phosphogluconate, induction of glucose-6-phosphate dehydrogenase in *E. coli* by superoxide is most interesting (20, 29). Given the probable importance of the Entner-Doudoroff pathway for aerobic growth on carbohydrates (1), we hypothesize that increased synthesis of glucose-6-phosphate dehydrogenase in the presence of superoxide leads to increased intracellular levels of 6-phosphogluconate. This would provide the organism with a mechanism for saturating 6-phosphogluconate dehydratase with the substrate to protect the enzyme from oxidative

TABLE 2. Putative promoter sequences within the *edd-eda* region

−35 region	Space class (bp)	−10 region	Size of intervening region (bp)	Start	Promoter
TTCACC	17	TACACT	5	A	P_1
TAAACT	18	TATTAT	3	A	P_2
CTGTCC	16	CAACCT	7	T	P_3
CCGAAC	16	CACTTT	13	A	P_4

stress under the conditions in which it would be most needed. The finding that superoxide dismutase mutants are impaired for growth on gluconate indicates that an increased 6-phosphogluconate level alone is not fully protective (16).

It has become apparent that the sequence of the *E. coli eda* gene has already been reported as part of a long-term study by Patil and Dekker (27). The enzyme purified by them from *E. coli*, KHG aldolase, is in fact identical to KDPG aldolase (37). They had previously commented on the similarity of this enzyme to KDPG aldolase from *P. putida*. A comparison of KDPG aldolase enzymes from three different organisms is shown in Fig. 3. Two active-site residues have been conserved in each of these enzymes. The role of KDPG aldolase in metabolism is complex. In *E. coli*, the enzyme is present at a branch point where catabolism of gluconate and hexuronic acids converge (14). Furthermore, there is substantial evidence that KDPG is a toxic metabolic intermediate, perhaps making it necessary to elevate KDPG aldolase to cope with flux through either pathway (14). The need for high basal levels of KDPG aldolase under conditions that do not result in KDPG formation is not so easily explained. It is tempting to speculate that KHG aldolase plays a role in regulating glyoxylate levels in *E. coli*. Gupta and Dekker (17) showed that glyoxylate could be mineralized to carbon dioxide via a pathway involving condensation with pyruvate, catalyzed by KHG aldolase, to form KHG. Low basal expression of isocitrate lyase could result in continuous synthesis of low levels of glyoxylate that might otherwise accumulate to toxic levels in the absence of the aldolase (21).

ACKNOWLEDGMENTS

Computer time was provided courtesy of the UNL Computer Resources Center. This work was supported by grants to T.C. from the U.S. Department of Energy (DE-FG02-90ER20002) and the Nebraska Energy Office (UNR-11) and to R.E.W. from the National Institutes of Health (Public Health Service grant GM27113).

We thank M. C. O'Neill for conducting computer-assisted promoter searches and Kenn Rudd for helpful discussions.

REFERENCES

1. **Adamowicz, M., T. Conway, and K. W. Nickerson.** 1991. Nutritional complementation of oxidative glucose metabolism in *Escherichia coli* via pyrroloquinoline quinone-dependent glucose dehydrogenase and the Entner-Doudoroff pathway. Appl. Environ. Microbiol. **57:**2012–2015.
2. **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.
3. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1520.
4. **Clarke, L., and J. Carbon.** 1976. A colony bank containing synthetic ColE1 hybrid plasmids representative of the entire *E. coli* genome. Cell **9:**91–99.
5. **Conway, T., R. Fliege, D. Jones-Kilpatrick, J. Liu, W. O. Barnell, and S. E. Egan.** 1991. Cloning, characterization, and expression 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.
6. **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.
7. **Conway, T., K. C. Yi, S. E. Egan, R. E. Wolf, and D. L. Rowley.** 1991. Location of the *zwf*, *edd*, and *eda* genes on the *Escherichia coli* physical map. J. Bacteriol. **173:**5247–5248.
8. **Cox, J. L., B. Cox, V. Fidanza, and D. Calhoun.** 1987. The complete nucleotide sequence of the *ilvGMEDA* cluster of *Escherichia coli* K-12. Gene **56:**185–198.
9. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
10. **Eisenberg, R. C., and W. J. Dobrogosz.** 1967. Gluconate metabolism in *Escherichia coli*. J. Bacteriol. **93:**941–949.
11. **Entner, N., and M. Doudoroff.** 1952. Glucose and gluconic acid oxidation of *Pseudomonas saccharophila*. J. Biol. Chem. **196:** 852–862.
12. **Fradkin, J. E., and D. G. Fraenkel.** 1971. 2-Keto-3-deoxygluconate-6-phosphate aldolase mutants of *Escherichia coli*. J. Bacteriol. **108:**1277–1283.
13. **Fraenkel, D., and S. Banerjee.** Deletion mapping of *zwf*, the gene for a constitutive enzyme, glucose-6-phosphate dehydrogenase, in *Escherichia coli*. Genetics **71:**481–489.
14. **Fraenkel, D. G.** 1987. Glycolysis, pentose phosphate pathway, and Entner-Doudoroff pathway, p. 142–150. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
15. **Fraenkel, D. G., and S. R. Levisohn.** 1967. Glucose and gluconate metabolism in an *Escherichia coli* mutant lacking phosphoglucose isomerase. J. Bacteriol. **93:**1571–1578.
16. **Gardner, P. R., and I. Fridovich.** 1991. Superoxide sensitivity of the *Escherichia coli* 6-phosphogluconate dehydratase. J. Biol. Chem. **266:**1478–1483.
17. **Gupta, S. C., and E. E. Dekker.** 1984. Malyl-CoA formation in the NAD^+-, CoASH-, and α-ketoglutarate dehydrogenase-dependent oxidation of 2-keto-4-hydroxyglutarate. J. Biol. Chem. **259:**10012–10019.
18. **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.
19. **Isturiz, T., E. Palmero, and J. Vitelli-Flores.** 1986. Mutations affecting gluconate catabolism in *Escherichia coli*. Genetic mapping of the locus for the thermosensitive gluconokinase. J. Gen. Microbiol. **132:**3209–3219.
20. **Kao, F. M., and H. M. Hassan.** 1985. Biochemical characterization of a paraquat-tolerant mutant of *Escherichia coli*. J. Biochem. **260:**10478–10481.
21. **Kornberg, H. L.** 1966. The role and control of the glyoxylate cycle in *Escherichia coli*. Biochem. J. **99:**1–11.
22. **Kovachevich, R., and W. A. Wood.** 1955. Carbohydrate metabolism by *Pseudomonas fluorescens*. III. Purification and properties of 6-phosphogluconate dehydratase. J. Biol. Chem. **213:** 745–756.
23. **Lewin, B.** 1990. Genes IV, p. 248–251. Oxford University Press, New York.
24. **Luria, S. E., and M. Delbruck.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28:**491–511.
25. **O'Neill, M. C.** 1989. *Escherichia coli* promoters. I. Consensus as it relates to spacing class, specificity, repeat substructure, and three-dimensional organization. J. Biol. Chem. **264:**5522–5530.
26. **O'Neill, M. C.** 1991. Training back-propagation neural networks to define and detect DNA-binding sites. Nucleic Acids Res. **19:**313–318.
27. **Patil, R. V., and E. E. Dekker.** 1992. Cloning, nucleotide sequence, overexpression, and inactivation of the *Escherichia coli* 2-keto-4-hydroxyglutarate aldolase gene. J. Bacteriol. **174:** 102–107.
28. **Pouyssegur, J. M., and F. R. Stoeber.** 1972. Rameau degradatif commun des hexuronates chez *Escherichia coli* K12. Eur. J. Biochem. **30:**479–494.
29. **Rowley, D. L., A. J. Pease, and R. E. Wolf, Jr.** 1991. Genetic and physical analyses of the growth rate-dependent regulation of *Escherichia coli zwf* expression. J. Bacteriol. **173:**4660–4667.
30. **Rowley, D. L., and R. E. Wolf.** 1991. Molecular characterization of the *Escherichia coli* K-12 *zwf* gene encoding glucose-6-phosphate dehydrogenase. J. Bacteriol. **173:**968–977.

31. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, 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. **Stormo, G. D.** 1986. Translation initiation, p. 195–224. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, Stoneham, Mass.
34. **Suzuki, N., and W. Wood.** 1980. Complete primary structure of 2-keto-3-deoxy-6-phosphogluconate aldolase. J. Biol. Chem. **255:**3427–3435.
35. **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.
36. **Vinopal, R. T., J. D. Hillman, H. Schulman, W. S. Reznikoff, and D. G. Fraenkel.** 1975. New phosphoglucose isomerase mutants of *Escherichia coli*. J. Bacteriol. **122:**1172–1174.
37. **Vlahos, C. J., and E. E. Dekker.** 1988. The complete amino acid sequence and identification of the active-site arginine peptide of *Escherichia coli* 2-keto-4-hydroxyglutarate aldolase. J. Biol. Chem. **263:**11683–11691.
38. **Wolf, R. E.** 1980. Integration of specialized transducing bacteriophage lambda *c*I857 *St*68 d*gnd his* by an unusual pathway promotes formation of deletions and generates a new translocatable element. J. Bacteriol. **142:**588–602.
39. **Yager, T. D., and P. H. von Hippel.** 1987. Transcript elongation and termination in *Escherichia coli*, p. 1241–1275. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
40. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and M13mp19 vectors. Gene **33:**103–119.
41. **Zlabotny, R., and D. G. Fraenkel.** 1967. Glucose and gluconate metabolism in a mutant of *Escherichia coli* lacking gluconate-6-phosphate dehydrase. J. Bacteriol. **93:**1579–1581.