Locating essential Escherichia coli genes by using mini-Tn10 transposons: the pdxJ operon.

H E Takiff, T Baker, T Copeland, S M Chen and D L Court

Locating Essential Escherichia coli Genes by Using Mini-Tn10 Transposons: the pdxJ Operon

HOWARD E. TAKIFF,† TERESA BAKER,1 TERRY COPELAND,2 SU-MIN CHEN,2 AND DONALD L. COURT1*

Molecular Control and Genetics Section, Laboratory of Chromosome Biology,1 and Protein Structure Group, Laboratory of Molecular Virology and Carcinogenesis,2 ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21702

Received 13 August 1991/Accepted 11 December 1991

The mini-Tn10 transposon (Δ16Δ17Tn10) confers tetracycline resistance. When inserted between a gene and its promoter, it blocks transcription and prevents expression of that gene. Tetracycline in the medium induces divergent transcription of the tetA and tetR genes within the transposon, and this transcription extends beyond the transposon in both directions into the bacterial genes. If the mini-Tn10 inserts between an essential bacterial gene and its promoter, the insertion mutation can cause conditional growth which is dependent on the presence of tetracycline. Two essential genes in adjacent operons of Escherichia coli have been detected by screening for tetracycline dependence among tetracycline-resistant insertion mutants. These essential genes are the era gene in the rnc operon and the dpj gene in the adjacent pdxJ operon. The pdxJ operon has not been described previously. It consists of two genes, pdxJ and dpj. Whereas the dpj gene is essential for E. coli growth in all media tested, pdxJ is not essential. The pdxJ gene encodes a protein required in the biosynthesis of pyridoxine (vitamin B6).

Conditional mutations facilitate the study of bacterial genes in two ways. They allow detection of physiological changes under nonpermissive conditions and provide the opportunity to select suppressor mutations. In addition to conditional lethal mutations of the thermosensitive type, conditional detects dependent on expression from heterologous promoters are also valuable for study. Chow and Berg (6) designed a transposition vector, TnStacl, to detect essential genes in Escherichia coli as conditionally lethal transposition mutations. Mutants generated depend on induction of the tac promoter from the transposon to express essential genes when their normal expression has been blocked by insertion of the transposon. We show here that a Tn10 derivative can also be used to isolate a class of conditionally lethal mutations.

We have defined three genes of the rnc operon: rnc, era, and recO (31). These genes encode the proteins RNaseIII, a double-strand-dependent endonuclease (7); Era, a GTP-binding protein (5, 21); and RecO, required for DNA repair and RecF pathway recombination (25). Of these, only era is essential for bacterial growth (14, 31). We previously isolated an E. coli mutant in which a mini-Tn10 (Δ16Δ17Tn10; 33) inserted in the leader sequence of the rnc operon caused the cell to become conditionally dependent on tetracycline for growth. This insertion mutation, mc-40, caused era expression to be defective in the absence of tetracycline presumably because the mini-Tn10 element was blocking normal transcription from the rnc promoter. We proposed that tetracycline, which induces the divergent promoters for the tetA and tetR genes of the mini-Tn10 transposon (13), allowed transcription of these genes to continue beyond the Tn10 element and into the era gene.

In this study, we have examined expression of lacZ fused to the rnc operon both with and without the rnc-40 mini-Tn10 insertion (A40). We have also isolated and examined several other mini-Tn10 insertion mutations in the region of the bacterial chromosome near the rnc operon. One of these additional mini-Tn10 insertions also confers a tetracycline-dependent (Tetr) phenotype. This new tetracycline-dependent insertion allowed us to define an operon adjacent to the rnc operon which also contains an essential gene.

MATERIALS AND METHODS

Materials. Restriction endonucleases and DNA-modifying enzymes were purchased from Bethesda Research Laboratories, Inc. (Gaithersburg, Md.); New England Biolabs, Inc. (Beverly, Mass.); and Boehringer Mannheim Biochemicals (Indianapolis, Ind.). Nick translation kits were purchased from Bethesda Research Laboratories or Boehringer Mannheim. DNA oligonucleotides were purchased from Midland Certified Reagent Co. (Midland, Tex.). The Sequenase kit was obtained from U.S. Biochemical Corp. (Cleveland, Ohio). Enzymes and kits were used as recommended by the manufacturers.

Ampicillin was used at a concentration of 50 μg/ml and tetracycline was used at a concentration of 12.5 μg/ml. For tetracycline-dependent (Tetr) strains, tetracycline was often used at concentrations of 1.0 to 2.5 μg/ml; the actual concentrations will be indicated for the experiments.

Bacterial and phage strains. Details of the E. coli K-12 strains and λ bacteriophage used are shown in Table 1. The genetic, bacteriological, and phage techniques used are described elsewhere, as are standard recipes for LB, TB, and MS6 media (23, 29). PI transductions were performed by using P1vir as described elsewhere (23), with the modification that for antibiotic selections, cells were spread on a nitrocellulose filter (BA85; Schleicher & Schuell, Inc., Keene, N.H.) on an LB plate. After 2 h of incubation, the filter was transferred to an LB-antibiotic plate.

Plasmids. Plasmids derived from pACS1 are shown in Fig.
1, and their construction is described below. pACS1 (31) contains a 4.3-kb EcoRI bacterial fragment inserted in the EcoRI site of pBR322 (Fig. 1). pACS21 is pACS1 deleted from the BamHI site indicated in Fig. 1 to the BamHI site in pBR322. pACS3 is pACS1 deleted from the SphI site in Fig. 1 to the SphI site in pBR322. pDLC140 is pACS1 deleted from the SphI site in Fig. 1 to the SphI site in pBR322. pDLC132 is pACS1 deleted for the AflIII fragment indicated in Fig. 1. pDLC141 is pDLC140 deleted for the AflIII fragment indicated. pDLC142 is pACS1 deleted for the NsiI fragment between the two NsiI sites shown in Fig. 2. pACS42 is pACS1 deleted for the KpnI fragment between the two KpnI sites shown in Fig. 2. All of the preceding plasmids were made by cutting with the designated enzyme, diluting the DNA, and joining the cut ends with T4 DNA ligase. Ampicillin-resistant transformants were selected. pDLC145 is pACS1 deleted for the 4-nucleotide segment AGCT within the SacI restriction site sequence (Fig. 2). This was done by cutting with SacI and then adding T4 DNA polymerase in the presence of all four deoxynucleotide triphosphates to trim the 3′ protruding ends before joining them with T4 DNA ligase. pDLC149 and pDLC150 are pACS42 and pDLC145 deleted for the AflIII fragment analogously to pDLC132. All of the constructs shown in Fig. 1 were verified by appropriate restriction digest analysis to ensure that the proper DNA fragment had been lost. Only pDLC145 failed to be recut by the original enzyme used to create each respective deletion.

LacZ fusion plasmids used in this work are described in the text and in the legends to figures that illustrate them. These operon and protein fusion plasmids were constructed with the lacZ vectors made by Simons et al. (30) to detect promoters and measure promoter strength. Promoterless vectors pRS414 and pRS591 (30) were used to construct the protein fusions; plasmids pRS415 and pRS528 (30) were used to construct the operon fusions.

### FIG. 1. The 4.3-kbp EcoRI fragment map of the mc and pdxJ operons and plasmids. The genes mc, era, recO, pdxJ, and dpj are indicated on the map below the open and filled rectangles that represent the sizes of the respective genes. The promoters p_mc and p_pdxJ and terminator t_pdxJ are shown. The lightly stippled areas at each end of the EcoRI map are bacterial DNA segments outside the mc and pdxJ operons. Restriction sites used to generate plasmids are shown above the genes or in Fig. 2. The set of rectangles above the map represents DNA of plasmids generated from a clone of this EcoRI fragment in pBR322. Plasmids are named at the right. For each plasmid indicated, the open rectangles represent DNA that is present, and the dashed line represents the region deleted. A40 and A8, indicated by triangles, represent the insertion sites for the mini-Tn10 insertion. See Materials and Methods for details of each plasmid construction.
NrdI

1

TCGCGAGAAGAGGGTTAATCGCAAGCCGTCTTTATCGCACATACGCAAGTTTACACGGAGACCTGCTTACCGGAGATTTCTGAC

-35

AluII

98

GCAGACAATCGGCGGCGGAAAGGTGTAAGCTGCAAGCTTCCGCAATGTCTCTGAGGTATCTCTGCGCTTACGCCGCTT

-10

A8 mini-Tn10

KpnI

BamHI

277

CAT ATC GCT GAC CGC AAC GGC GCC GGC CCT ACC GCT TAC CCG GAT CCG GTC CGG GCC GGC TTG ATT ACC GCA

12 His Ile Ala Thr Leu Arg Asn Ala Arg Gly Thr Thr Ala Pro Asp Pro Asp Pro Glu Ala Ala Phe Ile Ala Glu

361

Gln Ala Gly Ala Asp Gly Ile Thr Val His Leu Arg Glu Asp Arg Arg His Ile Thr Asp Arg Asp Val Ala Glu

421

ATC CTG CAG TCT AGT GAT ACC CCC ATT GCT CAT TTA CTT GAA GAT GCT GGC AAC ATT GCA

60 Ile Leu Arg Gin Thr Leu Thr Arg Met Asn Gin Ala Met Ala Val Thr Met Ala Val Ile Ala Ala

KpnI

BamHI

493

GTT GAG ACG AAG CCA CAT TGC TGC TCG GTA CCG GAA AAG GCT GAA GTA ACA ACC GAA GGC GCC CTC

84 Val Glu Thr Lys Pro His Phe Cys Cys Leu Val Pro Glu Lys Arg Glu Val Thr Thr Gly Leu

565

GAT GTC GCA GGC GCG CCT GCT AAG ATT GCG GAT GCC TGG AAA ATG CCT GCA GAT GGC GGG ATT CAG TTT CTC

108 Asp Val Ala Gly Gin Arg Asp Lys Met Asp Gin Ala Met Ala Val Cys Val Arg Ser Gin Ala Gly Glu Val Ser

637

CTG TTT ATT GNC GCC GTC GAT GAA GAG CAG ATC AAA GCT GCG GCA GAG GCT GGC GCC CGG TTT ATT ACC CAT

132 Leu Phe Ile Asp Ala Asp Glu Glu Gin Ile Lys Ala Ala Ala Glu Val Gly Ala Pro Phe Ile Glu His

709

ACC GGT TGC TCG GAT GCT GCC AAA ACT GAC GCC GAA CAG GAG GCA GCG GGT GCC GCT ATT CCC GACC TAC

156 Thr Gly Cys Tyr Asp Ile Asp Thr Ala Asp Glu Gin Ile Lys Ala Ala Ala Glu Glu Ala Ala Asp Ile Ala Ala

811

GTT CCC ATT GCC GCA GGC ATC CTG GCT AAG ATT GAC GCC GAG GCG GCC ATC GCC AAA GCG GGC

180 Thr Phe Ala Ser Leu Glu Leu Val Asn Ala Gin His Gly Leu Thr Tyr His Asn Val Ile Ala

925

GAC GGC GCC ATC TCT GAT AGT CAT GAA CAT ACC CAT GAC CCC GTC CTT TTA AGC GAT GCA TTA GAT GCA TCT

204 Ala Ala Ile Pro Glu Met His Glu Asn Ile Gin Gin Gin Arg Gin Met Gin Arg Glu Thr Met Ala

999

GAT ATT GGT GAG ATC GCT GCC ATC GAA GAC GGT ATC GCC GCA GGC GCC GGC GCC GTA TTA

9 Asp Asp Val Gin Glu Ile Ala Arg Ile Glu Ala Val Ile Ala Arg Ser Gin Asp Arg Leu Arg Ala Val

1071

ACC CAT AAC TGA TCT GGC CCC AAA GAG AAA CAC CAC CCC ACC ACC CCC CCG CTT

33 Ser Asp Asn Glu Trp Ala Ile Trp Lys Thr His Gin Glu Arg Val Arg Gin Phe Phe Phe

BamHI

1143

AAA GAA GCC GCA AAA GCG TTT GCC ACC GGG ATC AAT GAT CCT GGG TTT ATT AAA CAA TTT GAA GTA TTC

57 Lys Phe Ala Ala Ala Lys Ala Phe Gly Thr Gly Ile Arg Asn Glu Leu Ala Phe Asn Gin Glu Phe Phe

1215

AAC GAT GAG CTC GGC CCC AAA GAG ACC ACC ACC CCC ACC ACC CCC CCG CTT

81 Asp Asn Glu Leu Gly Lys Pro Arg Leu Arg Leu Trp Gly Leu Ala Leu Gly Leu Leu Gly Leu Gly Val

SauI

1287

GAA ATG CAT GCA ACG TCT GCA GAT GAG CAC GCC TAT GCT TGG CCC ACG GTA ATT ATT GAT GAA TAAAAAT

105 Ala Asn Met His Val Val Gin Gin Glu Arg Gin His Tyr Ala Cys Ala Thr Val Ile Ile Gin Ser

1361

ATACGTTCTTATTTTTAGAAAAACCACCGGATGGTAAAGGAAAGGCGGTTTATTTTTTTAAATTTGTGCGGACCC

FIG. 2. Nucleotide and deduced amino acid sequences of pdxA and dpj. Nucleotides are numbered from 1 to 1444. Relative to the maps in Fig. 1, nucleotide 1 corresponds to nucleotide position 2676 from the left EcoRI site, and nucleotide 1444 corresponds to nucleotide position 4119 from the left EcoRI site. The recognition sites for restriction enzymes are in boldface letters. Amino acids of the gene products are numbered and indicated below their codons. pdxA translation starts with the AUG at nucleotide position 244, and we assume that dpj starts at the AUG at nucleotide position 975. The two tandem translation stop codons UGA UGA of the recO gene are indicated by shaded letters upstream of the pdxA initiation codon. The pdx promoter is indicated by its −35 and −10 sites. Other pdx genes (pdxA and pdxB) each contain a sequence in the translation initiation region that is indicated by boldface italics (26). The target sequence of 9 nucleotides where the A8 insertion occurred is designated. The inverted repeat of a proposed rho-independent terminator is shown by inverted arrows. There is one discrepancy between this sequence and one published for the recO gene (25): at position 315 in our sequence, Morrison et al. found two Cs, whereas we find three Cs.
**Insertions and deletions**. Phage clone λTD1 was grown on the defective Tn10 transposon Δ16Δ17Tn10 (ΔTn10) donor strain RB132. Approximately 1 in 100 phage from such infections of RB132 carries a stable ΔTn10 insertion in its DNA. These rare phage with ΔTn10 inserts were not detected directly; instead, the population of phage was used to transduce E. coli to tetracycline resistance as described previously (31).

Briefly, to transfer the ΔTn10 from the phage to the chromosome of E. coli, an overnight culture of the nonlysogenic strain N4903 was infected with the population of λTD1 phage prepared on RB132. To reduce phage viability and thereby reduce cell killing, the phage for this infection were preirradiated with UV light (approximately 6 × 10³ ergs/mm²; Blak-Ray J-225 shortwave UV meter). Infected cultures of N4903 were grown on LB-tetracycline plates to select Tetr colonies in which the ΔTn10 was recombined to the chromosome. Simple lysogens of λTD1 cannot be formed because the phage is defective for the λ repressor.

For sequencing of the A8 insertion, the mutation in the chromosome was recombinated into λTD1lam21 and then subcloned to pUC19 as previously described (31).

**Physical mapping of the ΔTn10 insertions**. Bacterial DNA was isolated as described elsewhere (29). Southern blots of the bacterial DNA digested with various restriction endonucleases were hybridized to nick-translated phage or plasmid probes containing DNA from pACS1. The ΔTn10 insertional mapped by using BamHI, EcoRI, HindIII, and PstI digests of chromosomal DNA from the insertion strains.

**DNA sequencing**. Sequencing was performed with the Sequenase kit (U.S. Biochemical Corp.) and [y-32P]dATP using minipreparation plasmid DNA. Precise location of the mini-Tn10 insertion A8 was determined by sequencing from synthetic oligonucleotides (5'-GAAGGAACGTCAATTCCC-3') hybridized within the left and right ends, respectively, of the mini-Tn10 element cloned in pUC19, thereby priming synthesis to the flanking E. coli DNA (31). The DNA sequences of pdaX and dpj were determined from both strands of plasmid pACS1 by using a series of synthetic oligonucleotide primers. DNA sequence analysis was carried out with programs supplied by the University of Wisconsin.

**Expression and purification of 27-kDa protein**. TAP56 (pZSS1) was grown in LB medium (50 ml) containing ampicillin at 32°C to an optical density at 650 nm of 0.5 and transferred to 42°C for 2 h. Cells were harvested and lysed, and crude extracts were prepared as described by Chen et al. (5). The ~27-kDa protein (p27) was localized by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) analysis to the supernatant fraction after low-speed centrifugation.

A 450-μl portion of the supernatant was added to an equal volume of saturated guanidine hydrochloride, and the solution was then brought to pH 2 with the addition of 20% trifluoroacetic acid (TFA). The slightly turbid mixture was added to a Waters C-18 reverse-phase column in an LKB high-pressure liquid chromatography system. The column was washed with an equilibration buffer of 0.5 M ammonium bicarbonate pH 8.0, and then Trafic acid was added to 0 to 60% gradient of 0.05% TFA in acetonitrile at 1 ml/min over 60 min. Fractions were collected manually. Aliquots of protein peaks of interest were lyophilized and used for SDS-PAGE, amino acid composition, and amino-acid-sequence analysis.

**Amino acid composition analysis**. Protein from aliquots containing the 27-kDa protein (p27) were hydrolyzed with 6 N HCl at 110°C for 24, 48, and 72 h in vacuo in glass ignition tubes. The samples were then analyzed on a Beckman model 6300 amino acid analyzer.

**Amino-terminal-sequence analysis**. The protein p27 was subjected to Edman degradation in an Applied Biosystems model 470A gas-phase Sequenator, and phenylthiohydantoin derivatives of amino acids were identified and quantitated on-line with the model 120A analyzer.

**Tryptic digestion**. Protein was dissolved in 0.5 ml of 0.1 M NH₄HCO₃ (pH 8), and trypsin was added to give a protein/trypsin ratio of 100:1. The reaction was allowed to proceed for 24 h at 37°C. The reaction was stopped by the addition of aqueous TFA to give a pH of 2. The mixture was then applied to a Waters C-18 column, and a 0 to 60% gradient of 0.05% TFA in acetonitrile developed over 2 h. Peptide peaks were collected manually, and aliquots were taken for amino acid analysis.

**Galactokinase measurements**. The galK termination vector used in this work was pMZ240, which is derived from pKF1800 described by McKenney et al. (22) by a modification made by Zuber et al. (34) in which the gal promoter on pKF1800 was replaced by the lac promoter. Conditions for cell growth and measurement of galactokinase units have been described elsewhere (34).

**Nucleotide sequence accession number**. The nucleotide sequence reported here has been submitted to GenBank and has received the accession number M76470.

**RESULTS**

**Localized mini-Tn10 transposon mutagenesis**. Phage λTD1, which contains the rnc operon in a 14.5-kbp DNA segment of the E. coli chromosome, was mutagenized with the mini-Tn10 transposon Δ16Δ17Tn10 (see Materials and Methods; 31). Briefly, λTD1 was used to infect strain RB132, which contains the mini-Tn10 transposon encoding tetracycline resistance. RB132 also expresses transposase at high levels. The lysate from such an infection contained many phage which carried mini-Tn10. This lysate was used to infect strain N4903 in order to transfer the mini-Tn10 insertions in the bacterial DNA segment of the phage to the bacterial chromosome by homologous recombination. The recombinant bacteria form tetracycline-resistant (Tet') colonies. We found two types of Tet' colonies: those that grew with or without tetracycline and those that grew only when tetracycline was present. The latter we have called tetracycline dependent (Tet'^d). Concentrations of tetracycline as low as 1.0 μg/ml can satisfy the growth requirement of the Tet'^d colonies, and a detoxified form of chlorotetracycline can be used to replace tetracycline (31). We have studied two of these Tet'^d mutants in detail. For our initial discussions, the insertion mutations will be referred to as A40 and A8.

**Control of rnc operon expression by the A40 insertion**. The A40 mutation is located in the leader sequence of the rnc operon (Fig. 1). In the absence of tetracycline, this insertion prevents expression of the era gene, which is essential for E. coli growth (31). To determine how expression of era is controlled by A40, we fused lacZ to the era gene and measured levels of β-galactosidase. Without the A40 insertion, β-galactosidase expression was dependent on the promoter for the rnc operon (Table 2, line 1 versus line 2). The presence of the A40 insertion blocked lacZ transcription from this rnc promoter (Table 1, line 1 versus line 3). In the absence of tetracycline, the expression of β-galactosidase in the insertion mutants was increased severalfold (Table 2, lines 3 and 4). This induction occurred even when the rnc...
TABLE 2. Effect of mini-Tn10 insertion A40 on expression of era-lacZ fusions

<table>
<thead>
<tr>
<th>Protein fusion</th>
<th>Amt (U) of β-galactosidase</th>
<th>Plasmid</th>
<th>Amt (U) of β-galactosidase</th>
<th>Plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pCF14</td>
<td>2,683</td>
<td>ND</td>
<td>pCF140</td>
<td>6,240</td>
</tr>
<tr>
<td>pCF22</td>
<td>29</td>
<td>ND</td>
<td>pCF220</td>
<td>242</td>
</tr>
<tr>
<td>pCF03</td>
<td>38</td>
<td>133</td>
<td>pCF30</td>
<td>82</td>
</tr>
<tr>
<td>pCF01</td>
<td>15</td>
<td>92</td>
<td>pCF010</td>
<td>250</td>
</tr>
<tr>
<td>pRS591</td>
<td>&lt;10</td>
<td>&lt;10</td>
<td>pRS528</td>
<td>58</td>
</tr>
</tbody>
</table>

* The indicated lacZ fusion plasmids are described in Fig. 3. TAP114 carrying the indicated plasmids was grown to 2 × 10^8/ml in LB medium plus ampicillin either with 2.5 μg of tetracycline per ml (+Tet) or without tetracycline (−Tet). ND, not determined. Tetracycline has no effect on expression of β-galactosidase from pCF14 or pCF140 in other strains containing a copy of Tn10 on the bacterial chromosome (data not shown).

promoter was deleted. Thus, when the A40 insertion is present, transcription of era derives only from the mini-Tn10 and is tetracycline dependent. However, this level of tetracycline-dependent expression from the insertion is still considerably less than that from the mc promoter.

**Defects caused by the A8 insertion.** Restriction digests and Southern blot hybridization analyses mapped the A40 insertion to the leader sequence of the mc operon (31) and the A8 insertion to near the end of the recO gene (data not shown). In the mutant with the A40 insertion, the divergent tetR and tetA genes present in the element are arranged with tetR in the same orientation as mc and era. In the mutants with the A8 insertion, the tetR and tetA genes are inverted relative to A40 (data not shown). Since both the tetA and tetR promoters are activated by tetracycline, the Tet^ phenotype of the A8 mutant strain may reasonably be explained by a mechanism analogous to that of the A40 mutant strain in which the tetR promoter is thought to direct era expression. Therefore, we postulated that an essential gene is positioned beyond A8 relative to recO and is dependent on transcription from the tetA promoter for its expression.

The A8 and A40 insertions in each require tetracycline for growth. Since pACS1 complemented Tet^ of both A40 and A8 strains (Fig. 1 and Table 3), we tested several other deletion derivatives of pACS1 for complementation. As determined previously (31), the Tet^ defect caused by A40 was complemented by plasmids carrying an intact era gene. The A8 mutation was complemented for Tet^ by all the plasmids in which the entire distal reading frame (dpr) in Fig. 1 downstream of the A8 insertion (pDLC140, pDLC141, pACS42) was intact. In other words, a strain with the A8 insertion could grow without tetracycline if it contained a plasmid with an intact copy of the second open reading frame. Thus, the distal gene, like era, is vital for *E. coli* growth in all media that we have tested.

Unlike the strain with the A40 insertion, the strain with the A8 insertion failed to grow even in the presence of tetracycline on M56 minimal medium supplemented with glucose or any of several other carbon sources tested (data not shown). Casamino Acids did not satisfy the growth requirement, but a vitamin mixture permitted growth. Further testing indicated that pyridoxine was required for growth.

The plasmids described above to test for complementation of the pyridoxine requirement, i.e., for growth on minimal medium with tetracycline. The A8 defect on minimal plates with tetracycline present was complemented by those plasmids containing the first gene, i.e., pdsU in Fig. 1 (pDLC140, pDLC141, pDLC145). Therefore, the disruption of this gene by the A8 insertion causes the pyridoxine requirement.

We had also noticed that λ formed clear plaques on lawns of strains with the A8 insertion. When pyridoxine or yeast extract was added to the TB-tetracycline plates used for λ plaque tests, λ formed turbid plaques on the A8 insertion.
TABLE 4. Complementation of pdx mutants

<table>
<thead>
<tr>
<th>Strain</th>
<th>Allele*</th>
<th>Growth with plasmids**: None</th>
<th>pDLC132</th>
<th>pDLC149</th>
<th>pDLC150</th>
</tr>
</thead>
<tbody>
<tr>
<td>WG1</td>
<td>pdx⁺</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>WG1027</td>
<td>pdxJ151</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>WG1050</td>
<td>pdxJ161</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>WG1123</td>
<td>pdxJ178</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>WG1129</td>
<td>pdxK180</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WG1128</td>
<td>pdxJ186</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>WG1135</td>
<td>pdxK191</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>HT253</td>
<td>pdxJ8::ΔTn10</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

* Strains were plated for isolated colonies on M56 glucose with Casamino Acids added. All strains grew on this medium when supplied with pyridoxine. Strains WGI(pdxJ), WG3(pdxJ8), WG25(pdxA25), WG37(pdxJ473), WG1145 (pdxJ), and WG1468(pdxB) were also tested and failed to be complemented by pDLC132. All WG strains are E. coli B derivatives of WG1, a prototrophic strain (9).

** +, complete complementation for growth as with pyridoxine added; -, no or very reduced growth relative to the pyridoxine control. All plasmids are deleted for the AflII fragment and part of the pdx promoter, but enough residual promoter activity remains to complement at high copy number (Table 6, plasmid pDLC135). The AflII deletion removes an EcoB restriction site of pAC51.

Thus, the clear-plaque phenotype observed was most likely caused by a deficiency of pyridoxine in TB medium which resulted in poor cell growth. Poor cell growth is known to reduce the turbidity of λ plaques (20).

The pdxJ and pdxK genes were previously mapped at min 55 on the E. coli chromosome (1, 2, 9) very close to rnc. These mutants require pyridoxine for growth. Six strains with pdxJ or pdxK point mutations (9) could be complemented for growth in the absence of pyridoxine by pDLC132 and pDLC150 (Table 4). Other pdx genes, located at other sites on the chromosome, are not complemented by these plasmids (Table 4). Therefore, the pdxJ and pdxK mutations are likely to represent different alleles of the same gene that differ only in their leakiness, as suggested previously (9). In keeping with the genome map (1, 2), we have named the gene inactivated by A8 pdxJ. We have named the distal gene dpj, since it is downstream of pdxJ in the operon. We have no results that indicate that dpj is involved in pyridoxine biosynthesis.

**Nucleotide sequence of the pdxJ operon and insertion site of A8.** Figure 2 shows the nucleotide sequence that we determined for the region from an NruI site in recO to a SphI site near the right end of the EcoRI fragment (Fig. 1). Both DNA strands were sequenced by using several synthetic oligonucleotides to prime dideoxy-chain termination reactions. Analysis of the sequence revealed two adjacent open reading frames that nearly span the region from beyond recO to the SphI site (Fig. 2). Both open reading frames use the same template DNA strand that encodes the recO operon genes, and the initiating methionine codon for the first reading frame begins 9 bp beyond recO. We believe that this is the pdxJ gene and that the open reading frame in which the SacI site lies is dpj.

The pdxJ gene potentially encodes a protein of 243 amino acids (26,384 Da), and the dpj gene encodes a protein of 126 amino acids (14,052 Da). From their amino acid compositions, isoelectric points of 5.9 and 10.0 were predicted, respectively. Both genes use rare codons characteristic of other genes in E. coli that are expressed at low levels, including pdxJ and pdxK (26, 28).

We determined the precise location of the A8 insertion by sequencing a subclone in pUC19 (see Materials and Methods). The insertion site is 36 bp beyond the 3' end of recO and within the pdxJ reading frame (Fig. 2). The insertion target sequence TGTAGGGCC was 9 nucleotides in length and was repeated at each border of the mini-Tn10, as expected for Tn10 transposition (15).
TABLE 5. Amino acid composition of p27

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Found*</th>
<th>Predicted*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cysteine</td>
<td>ND</td>
<td>4</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>23</td>
<td>16</td>
</tr>
<tr>
<td>Asparagine</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Threonine</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Serine</td>
<td>70</td>
<td>20</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>30</td>
<td>9</td>
</tr>
<tr>
<td>Glutamine</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Proline</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Glycine</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Alanine</td>
<td>31</td>
<td>40</td>
</tr>
<tr>
<td>Valine</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Methionine</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>Leucine</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Histidine</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Lysine</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>Arginine</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>ND</td>
<td>0</td>
</tr>
</tbody>
</table>

* Based on 24-, 48-, and 72-h hydrolys. ND, not determined.

* From Fig. 2.

* The amino-terminal methionine has not been included.

TABLE 6. Mapping the pdxJ operon promoter with pdxJ-lacZ fusions

<table>
<thead>
<tr>
<th>Gene fusion</th>
<th>Amt (U) of β-galactosidase</th>
<th>Operon fusion</th>
<th>Amt (U) of β-galactosidase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid</td>
<td></td>
<td>Plasmid</td>
<td></td>
</tr>
<tr>
<td>pDLC129</td>
<td>2,838</td>
<td>pDLC130</td>
<td>5,544</td>
</tr>
<tr>
<td>pDLC138</td>
<td>2,742</td>
<td>pDLC139</td>
<td>3,077</td>
</tr>
<tr>
<td>pDLC134</td>
<td>113</td>
<td>pDLC135</td>
<td>431</td>
</tr>
<tr>
<td>pDLC147</td>
<td>58</td>
<td>pDLC148</td>
<td>259</td>
</tr>
<tr>
<td>pDLC144</td>
<td>&lt;10</td>
<td>pRS415</td>
<td>23</td>
</tr>
</tbody>
</table>

* The lacZ fusion plasmids are described in detail in Fig. 6. TAP56 carrying the respective plasmids was grown overnight in LB medium plus ampicillin. Fresh cultures were inoculated in the same medium at a 1:100 dilution and grown to 2 × 10^6 to 4 × 10^6 cells per ml.

Expression of the pdxJ operon from the λ pL promoter.

Previously, we had made fusions of the pL promoter of phage λ to the rnc and era genes and expressed the products of these genes at high levels (5). In similar experiments, we have looked for pL-promoted expression of the genes beyond recO. pCE21 and pZS21 (Fig. 4) produced a 27-kDa protein (p27) from the pL promoter but only in cells containing λ N protein; i.e., no 27-kDa protein was produced in the

FIG. 6. Defining the pdxJ promoter by gene and operon lacZ fusions at the BamHI site in pdxJ. (A) In construct 1, the BamHI fragment (open rectangle) was isolated and inserted in the orientation shown at the BamHI site of the lacZ vectors pRS414 and pRS415 (30), creating pDLC128 and pDLC130, respectively. In construct 3, the EcoRI-to-BamHI fragment from pDLC132 (Fig. 1), was used to generate pDLC134 and pDLC135 by replacing the EcoRI-to-BamHI fragments from pRS414 and pRS415, respectively. A 19-bp synthetic DNA, CGCTTACCCGATGTGGC, was created from complementary oligonucleotides. This synthetic DNA had single-stranded EcoRI- and AflII-compatible ends. It was used to replace the EcoRI-to-AflII fragments of pDLC134 and pDLC135, creating pDLC138 and pDLC139. pDLC147 and pDLC148 are pDLC138 and pDLC139 but with the AflII site cut and then repaired with Klenow enzyme to add the 4 nucleotides TTAA (shown as an X). The sequence of each of these regions at the AflII site is shown in panel B. (B) The numbered sequences represent each construct shown in panel A. Line 1 is entirely wild-type sequence for this region. Lines 2 and 3 represent deletion derivatives in which wild-type sequence is deleted from the left. The underlines indicate nucleotide sequences changed relative to the wild-type sequence. Line 4 represents the same plasmid and sequence as in line 2 except that the AflII site between the −35 and −10 regions has been cut and then filled in with Klenow enzyme to add 4 nucleotides (TTAA). This shifts the upstream sequence relative to the promoter. Below the sequences is the consensus sequence for the −35 and −10 regions of a promoter with the optimal 17-nucleotide spacing (27). The most-conserved and most-important nucleotides in the −35 and −10 hexamers are in capital letters, and the less-conserved nucleotides are in lowercase letters. Within each of the four sequences, nucleotides in the appropriate −35 and −10 regions that match the consensus are shown in boldface type. The arrow seven nucleotides beyond the −10 region represents the proposed start and direction of transcription for the promoter. Promoter activity is normalized relative to the 100% β-galactosidase values for the protein (pDLC129) and operon (pDLC130) fusions, which are 2,838 and 5,544 U, respectively.
N mutant strain, TAP106 (Fig. 5). This N effect will be analyzed in the Discussion. pCE22 failed to express the 27-kDa protein (data not shown). pCE22 is deleted for both \( \text{pdxJ} \) and \( \text{dpj} \). Since \( \text{pdxJ} \) can encode a 27-kDa protein, we considered it to be the source of the expressed protein. No other protein in the size range (~14 kDa) expected for \( \text{dpj} \)

was found in the expressed cells.

We expressed the 27-kDa protein in TAP56 containing p2SZ1 and prepared crude extracts by freeze-thaw lysis of the cells. The 27-kDa protein, which made up more than 40% of the total cellular protein, was in the supernatant after a low-speed centrifugation. This supernatant was then fractionated by reverse-phase liquid chromatography, and the major protein peak (p27) was subjected to amino-terminal-sequence analyses as well as composition analysis. The amino-terminal sequence of 18 amino acids matched that deduced from the coding sequence except that methionine, the first amino acid residue in the predicted sequence, was absent in the protein. Two internal peptides were purified from a trypsin digest of the protein. Their amino acid sequences corresponded with the proposed protein sequence from positions 33 to 44 and from positions 134 to 146.

The amino acid composition of p27 is given in Table 5 and compared with the predicted amino acid composition. Overall, the agreement is quite good. Serine levels are high, but they could be due to contamination of free amino acid. The only discrepancy is that the level of alanine is somewhat low.

The \( \text{pdx} \) promoter and terminator. Protein and operon fusions of \( \text{pdxJ} \) to \( \text{lacZ} \) at the \( \text{BamHI} \) site in \( \text{pdxJ} \) were constructed to detect and map the promoter for this gene. Differences in \( \beta \)-galactosidase levels between different deletion mutants indicated that an active promoter was located at the \( \text{AflII} \) site just upstream of \( \text{pdxJ} \) (Fig. 6; Table 6). The \( \text{BamHI} \) fragment containing the entire \( \text{recO} \) gene was joined to \( \text{lacZ} \) in both protein (pDLC129) and operon (pDLC130) fusions. We note that neither of these fusions complemented the \( \text{recO} \) mutant strain HT210 for UV, indicating that the \( \text{recO} \) (\( \text{p}_{\text{recO}} \)) promoter had been deleted. However, both fusion plasmids yielded high levels of \( \beta \)-galactosidase activity, indicating the presence of a promoter on the \( \text{BamHI} \) fragment. Plasmids bearing only the \( \text{AflII} \)-to-\( \text{BamHI} \) fragment fused to \( \text{lacZ} \) (pDLC134 and pDLC135) produce greatly reduced levels of \( \beta \)-galactosidase. This indicates that the promoter for \( \text{pdxJ} \) lies in the interval between the \( \text{BamHI} \) and \( \text{AflII} \) sites missing in pDLC134 and pDLC135. High \( \beta \)-galactosidase levels were restored by replacing 19 bp of wild-type sequence upstream of the \( \text{AflII} \) site (pDLC138 and pDLC139). This result suggested that the upstream boundary (the ~35 region) of the promoter was within the replaced 19-bp region (Fig. 6B, line 2). If this 19 bp defines the ~35 region, then the \( \text{AflII} \) site is located in the 17-bp spacer between the ~35 and ~10 regions of the promoter. When the \( \text{AflII} \) site of pDLC138 or pDLC139 was mutated by insertion of 4 bp, the promoter became inactive (Table 6, line 4). We believe, therefore, that the ~35 and ~10 positions indicated in Fig. 3 and 6B define the promoter for the \( \text{pdxJ} \) operon.

We have also tested pDLC129 and pDLC130 in TAP56, TAP56 mc-14, and TAP56 pdxJ8 strains and found no difference in expression between these strains (data not shown). Thus, we see no regulatory control on \( \text{pdxJ} \) by the \( \text{mc} \) and \( \text{pdxJ} \) gene products.

A transcription terminator beyond \( \text{dpj} \). Between the end of \( \text{dpj} \) and the \( \text{Sppl} \) site is an extensive inverted repeat sequence shown in Fig. 2 by the underlined arrows. This site has characteristics of a rho-independent terminator, namely, an RNA loop between a GC-rich stem segment followed by a uridine stretch in the RNA. Since the inverted repeat in this region includes both the GC-rich stem and the adjacent uridine segment, the symmetric DNA sequence may be capable of terminating transcription in both orientations (11).

We placed a 150-bp DNA fragment (NstI to Sppl in Fig. 2) containing the inverted repeat in a transcription test plasmid between the lac promoter and a reporter \( \text{galK} \) gene encoding galactokinase. Without the fragment inserted, the vector produced 1,087 U of galactokinase. Twenty-one units and 90 U were made with the proposed terminator fragment inserted in the normal and inverted orientations, respectively. Thus, it is an effective terminator in both orientations.

DISCUSSION

We have demonstrated that in addition to its ability to transposes, the mini-Tn10 element \( \Delta 16 \text{A17Tn10} \) (33) has two properties that allow detection of essential genes in \( \text{E. coli} \). First, when the element inserts in a transcribed region, it causes a strong polarity effect and prevents gene expression from the natural promoter. Second, when tetracycline is present in the medium, the two divergent genes within the transposon are induced for expression (13) and the resulting transcriptions of \( \text{tetA} \) and \( \text{tetR} \) extend beyond both ends of the element into adjacent bacterial sequences. Thus, while blocking expression of downstream genes from the bacterial promoter, the mini-Tn10 can conditionally activate those downstream genes following induction with tetracycline. If the downstream genes are essential for growth, the mutants may grow conditionally upon the addition of tetracycline. In this manner, the mini-Tn10 insertions have allowed us to find two essential genes, \( \text{era} \) and \( \text{dpj} \). Although the normal \( \text{Tn10} \) transposable element causes polarity (16), there are no reports that its tetracycline-induced transcripts extend beyond its boundaries into bacterial genes as the mini-Tn10 transcripts do.

We suspect that not all genes will be expressed at functional levels from a mini-Tn10 insert, especially those genes which require high levels of expression. Both \( \text{era} \) and \( \text{dpj} \) contain rare codons characteristic of \( \text{E. coli} \) genes that are expressed at low levels. Our measurements of Era protein in the cell substantiated its low level (3). The level of tetracycline-dependent Era protein expressed from A40 is clearly less than that from the natural \( \text{mc} \) promoter (Table 2); however, it is still sufficient to allow cell growth. We note that transcription from the A40 element should include both \( \text{mc} \) and \( \text{era} \), since A40 is located upstream of both genes, yet the cells remain phenotypically \( \text{Rnc}^- \). We presume that the cell is making reduced levels of \( \text{RNase III} \) that are below a threshold required for an \( \text{Rnc}^+ \) phenotype. Thus, \( \text{mc} \) is an example of a gene that cannot be restored to a normal phenotype by transcription from the mini-Tn10. Still, we believe that the mini-Tn10 will be useful for detecting genes, both essential and nonessential, that are expressed at low levels.

Takiff et al. (31) isolated a mini-Tn10 insertion (A14) within the \( \text{mc} \) gene. This strain is not Tet. We have examined \( \text{era-lacZ} \) fusions of this A14 mutant and found that its expression is higher than that found for the strain with the A40 insertion and that expression is constitutive, i.e., not inducible by tetracycline. Sequence analysis of the junction with the mini-Tn10 and \( \text{mc} \) revealed that a weak promoter was probably created. It is composed of a poorly conserved ~35 element (TAGGGG) that is present in the inverted repeat at the ends of the mini-Tn10 and spaced 17 bases from

\begin{table}
\centering
\begin{tabular}{|c|c|c|}
\hline
Gene & Transcription & Termination
\hline
\text{dpj} & Yes & Yes
\hline
\text{pdxJ} & Yes & Yes
\hline
\text{recO} & Yes & Yes
\hline
\end{tabular}
\caption{Summary of \( \text{pdxJ} \) promoter and terminator characterization.}
\end{table}
A 10 sequence (TACACT) present in the rnc gene adjacent and just downstream of the insertion.

The gene pdxJ is inactivated by the mini-Tn10 insertion A8, and dpj is not expressed because of transcriptional polarity. Since dpj is essential for growth, cells with the A8 insertion grow only with tetracycline present or if carrying a plasmid that contains an intact dpj gene. The function of the dpj product remains unknown. The dpj gene is known only by its codons, amino acid sequence, and operon relationship with pdxJ. dpj encodes a highly basic protein of 14.0 kDa. Comparison of its amino acid sequence with those in the translated genetic data base failed to reveal any significant similarities with other gene products.

The pdxJ gene encodes a protein of 26.4 kDa that functions in the biosynthesis of pyridoxal phosphate. The A8 insertion interrupts this gene and causes a cellular requirement for pyridoxal or pyridoxine (8, 10). Lam et al. in the accompanying paper (17) suggest that PdxJ may take part in the synthesis of pyridoxine by the condensation of 1-deoxyxylulose and 4-hydroxyphenylene.

The mec operon consists of three genes: mec, era, and recO. Three observations support the proposal that transcription of recO is dependent on the mec promoter. First, insertion mutations in mec and era between rrmB and recO are polar on recO. These polar mutations cause the cell to become sensitive to UV irradiation, a phenotype of recO mutants (25, 31). The BamHI fragment that includes recO and is found in pDLC129 and pDLC130 (Fig. 6) does not complement the recO mutant strain HT210. This indicates that the recO promoter is upstream of Era. Complementation of recO seen previously with this BamHI fragment (25) may have been caused by transcription from plasmid promoters flanking the fragment. The lacZ plasmid vectors used here were constructed to minimize interference by endogenous plasmid promoters (30). Finally, transcription from the rnc promoter is reduced as it enters recO; only about 1 in 10 transcripts continues through the recO gene (3).

The pdxJ operon promoter is located within the recO gene, and the pdxJ ribosome-binding and initiation site overlaps with the translation termination codons of recO. We have not found a typical rho-dependent terminator (4) for rnc operon transcription in the 100 nucleotides of sequence beyond recO. However, we believe for the following reasons that there is a strong terminator, perhaps a rho- or other factor-dependent terminator, in this region. Transcription initiated from the powerful λP1 promoters in plasmids pCE21 and pZS15 produces the pdxJ protein. However, with the same P1 fusions, large amounts of PdxJ are made when the λN gene product is provided (Fig. 4). N causes RNA polymerases that initiate at P1 to become insensitive to terminators; i.e., N causes transcription anti-termination (9). We assume that the terminator that blocks transcription from P1 under N- conditions is at the natural end of the rnc operon within pdxJ. If this is correct, we are left with the problem of how the promoter located just upstream of pdxJ in recO can bypass this terminator. We speculate that the answer may lie in the proximity of the pdxJ promoter to the proposed transcription terminator, i.e., that it may be too close. For rho-dependent termination to be effective, the terminator must be at least 100 nucleotides from the promoter (24, 34).

ACKNOWLEDGMENTS

This research was supported by the National Cancer Institute under contract NO1-CO-74101 with A.B.L. We also acknowledge support from the National Science Foundation U.S.-Japan Cooperative Science Program under grant INT-8815935.

We thank Zhao-Shan Zhang for constructing and characterizing pZS15 and Nina Costantino for providing several enzyme assays. Malcolm Winkler and Walter Dempsey provided us with unpublished information and strains. We are grateful for several helpful discussions with Stanley Brown and for the typing of the manuscript by Carolyn Redmond. We also acknowledge the help of Yoshikazu Nakamura in this work.

REFERENCES


