JOURNAL OF BACTERIOLOGY, Jan. 1989, p. 172–183
0021-9193/89/010172-12$02.00/0

Vol. 171, No. 1

Sequence of the *Pseudomonas aeruginosa trpI* Activator Gene and Relatedness of *trpI* to Other Procaryotic Regulatory Genes

MING CHANG, AYELE HADERO,† AND IRVING P. CRAWFORD*
Department of Microbiology, University of Iowa, Iowa City, Iowa 52242

Received 17 June 1988/Accepted 4 October 1988

In *Pseudomonas aeruginosa*, the *trpI* gene product regulates the expression of the *trpBA* gene pair encoding tryptophan synthase. *trpI* and *trpBA* are transcribed divergently. The *trpI* DNA sequence and deduced amino acid sequence were determined. The *trpI* start codon was found to be 103 base pairs from that of *trpB*. *trpI* encodes a 293-residue protein and the size of the *trpI* gene product, measured on sodium dodecyl sulfate-polyacrylamide gels, was close to that calculated from the amino acid sequence. The amino acid sequence of *trpI* resembles that of *Enterobacter cloacae ampR*, the regulatory gene for the *ampC* cephalosporinase. The N-terminal portions of *trpI* and *ampR* resemble corresponding portions of *ilvY*, *metR*, and *lysR* in *Escherichia coli* and *nodD* in *Rhizobium meliloti*. This resemblance may help to define a *trpI*-related family of activator proteins sharing a common structural plan.

In *Pseudomonas aeruginosa*, seven *trp* structural genes, scattered at four chromosomal locations, accomplish the synthesis of tryptophan from its precursor, chorismate (Fig. 1) (6, 11, 27). Early studies showed that in fluorescent pseudomonads the genes for tryptophan synthase, *trpB* and *trpA*, are coordinately induced by the substrate, indoleglycerol phosphate, rather than repressed by the product, L-tryptophan (6, 13). More recently, the *P. aeruginosa* PAC174 *trpBA* genes along with their control region were subcloned from an R68.44-derived R-prime plasmid containing about 114 kilobases (kb) of *P. aeruginosa* chromosomal DNA into plasmid pBR322 (19, 27). The resulting 8.5-kb plasmid, pZAZ167, enables an *Escherichia coli trpE trpB trpA* auxotroph to grow on anthranilate in place of tryptophan (19, 27). Several deletion mutants of pZAZ167 located upstream of the *trpBA* genes cause low, constitutive levels of tryptophan synthase even in the presence of the inducer indoleglycerol phosphate but can be restored to normal inducibility in *trans* (28). Therefore, it was proposed that a *trpI* gene, mapping upstream of the *trpBA* genes, encodes an activator-like mediator responsible for the induction of the *trpBA* genes (28). In this paper, we establish the location, orientation, and size of the *trpI* gene and present its DNA sequence and deduced amino acid sequence. Comparison of the *trpI* amino acid sequence with that of *ampR* in *Enterobacter cloacae*, *ilvY*, *metR*, and *lysR* in *E. coli*, and *nodD* in *Rhizobium meliloti* shows that these activator proteins have significant similarity in their N-terminal portions. The diversity in their C-terminal portions may reflect differences in the inducer molecules with which they interact.

MATERIALS AND METHODS

Bacterial strains, plasmids, and media. The bacterial strains and plasmids used or constructed in this study are shown in Table 1. The bacterial strains were cultured in Vogel-Bonner minimal medium E (38), in LB medium, or on X-Gal (5-bromo-4-chloro-3-indoxyl-β-D-galactopyranoside) agar plates (37). X-Gal was obtained from Sigma Chemical Co., St. Louis, Mo. To ensure retention of vectors or recombinant plasmids, appropriate antibiotics were added in the following amounts (micrograms per milliliter): ampicillin, 200; tetracycline, 10; and chloramphenicol, 12.5. X-Gal plates were used to screen pUC plasmid derivatives transformed into JM83 or JM101 cells. Vogel-Bonner minimal medium was used for growth rate tests.

Growth rate. Growth rate tests were used to assess the activity of the *trpI* gene product. The principle is based on the observation of Manch and Crawford (28) that *E. coli* IC1107 harboring an uninducible deletion plasmid such as pZAZ131 grows slowly on minimal agar plates supplemented with anthranilate or indole but grows well on tryptophan. However, cotransformation of pZAZ131 with a *trpI*-carrying plasmid such as pZAZ202 allows *E. coli* IC1107 to grow as rapidly on anthranilate as on tryptophan, but it still grows slowly on indole.

In this study, growth rate tests were carried out by cotransformation of two compatible plasmids, one carrying the *trpI* gene and the other carrying the *trpBA* gene pair, into *E. coli* IC1107. The transformed cells were streaked on four kinds of minimal plates. After 12 to 36 h of incubation at 37°C, the extent of growth of the transformed cells was estimated by the size of isolated colonies. The four minimal plates used were labeled MAHC, MTAHC, MAAHC, and MIAHC. MAHC was a Vogel-Bonner minimal agar plate containing 0.05% (wt/vol) acid-hydrolyzed casein and 0.2% glucose but no other supplements. MTAHC, MAAHC, and MIAHC were MAHC containing 0.1 mg of tryptophan, anthranilate, and indole per plate, respectively. The MAHC plate was used to test for contamination of the transformed cell culture, since the IC1107 transformed cells remain *trpE* auxotrophs. The MIAHC plate was used to detect the occurrence of any constitutive mutants.

A *trpI*-containing plasmid was defined as $TrpI^+$ when the transformed cells grew as fast on the MAAHC plate as on the MTAHC plate. If the transformed cells grew slowly on the MAAHC plate, the plasmid was defined as $TrpI^-$. Cells transformed with either a $TrpI^+B^+A^+$ plasmid or a $TrpI^-B^+A^+$ plasmid were always used as controls in the growth rate test.

Plasmid DNA manipulation. The procedures for plasmid purification, restriction enzyme digestion, DNA ligation, agarose gel electrophoresis, and polyacrylamide gel electro-

* Corresponding author.
† Present address: Department of Pediatrics, SUNY Health Sciences Center at Brooklyn, Brooklyn, NY 11203.

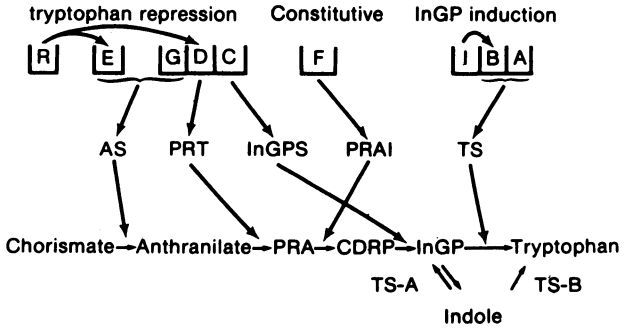


FIG. 1. Tryptophan biosynthetic pathway, with the corresponding *trp* genes in *P. aeruginosa*. The *trpR* gene encodes a repressor acting on *trpE* and the *trpGDC* cluster. The *trpI* gene encodes an activator inducing the *trpBA* gene pair. The *trpE* and *trpG* genes encode the large and the small subunits of anthranilate synthase, and *trpA* and *trpB* encode the α and β subunits of tryptophan synthase. The gene products and their participation in biosynthetic reactions are indicated by arrows. Abbreviations: AS, anthranilate synthase; PRT, anthranilate phosphoribosyltransferase; InGPS, indoleglycerol phosphate synthase; PRAI, *N*-phosphoribosylanthranilate isomerase; TS, tryptophan synthase; PRA, *N*-phosphoribosylanthranilate; CDRP, 1-(*o*-carboxyphenylamino)-1-deoxyribulose phosphate; InGP, indoleglycerol phosphate; TS-A, tryptophan synthase A reaction; TS-B, tryptophan synthase B reaction.

phoresis were those described by Maniatis et al. (29). When pACYC184-derived plasmids were purified, 300 mg of spectinomycin was added per liter of culture in the late exponential phase to amplify plasmids. *E. coli* cells were made competent for transformation by the $CaCl_2$ method (10). Plasmid pMI10 was pZAZ125 deleted of two *Ava*I fragments. Plasmid pZAZ125 has a *Bam*HI-*Sal*I fragment (Fig. 2A) inserted into the *tet* gene of pACYC184. The only *Ava*I site in pACYC184 is located downstream of the *Sal*I site and is within the *tet* gene. This *Ava*I site and two other *Ava*I sites in the *trpI* gene (Fig. 2B) of pZAZ125 were cleaved; the rest of plasmid was then self-ligated, generating pMI10. Plasmid pMI110 was pZAZ167 with its *Sph*I-*Sph*I fragments removed (Fig. 2A). Plasmid pBD6 was pMIT3 deleted of *Bal*I-*Bal*I and *Bal*I-*Stu*I fragments in the middle of the *trpI* gene (Fig. 2B).

BAL 31 exonuclease was used to decrease the size of DNA fragments or generate various-sized deletions. The experimental procedure used is described by Maniatis et al. (29). Each batch of plasmid DNA was digested for various times by a fixed amount of enzyme to decide the proper conditions. The enzyme used was obtained from Bethesda Research Laboratories, Inc., Gaithersburg, Md. The buffer and amount of enzyme used were those suggested by the manufacturer.

Plasmids pUC18 and pUC19 were often used as cloning vectors because their makeup facilitates identification of recombinants (41). The multiple cloning sites in pUC18 and pUC19 are positioned in opposite directions relative to the *lac* promoter. A *Sal*I-*Pst*I fragment containing the entire *trpI* region was subcloned into the *Sal*I and *Pst*I sites of pUC19 and pUC18, generating pMI40 and pMI50, respectively.

For DNA sequencing, DNA fragments were first labeled with ^{32}P by one of two methods: 5′ labeling with terminal transferase and [α-^{32}P]2′,3′-dideoxyadenosine triphosphate (42) or 3′ fill-in labeling with the large fragment of *E. coli* DNA polymerase I and the appropriate radioactive deoxynucleotide triphosphate (26). Sequencing reactions were carried out by the procedure of Maxam and Gilbert (26) and

TABLE 1. Bacterial strains and plasmids

E. coli strain or plasmid	Relevant characteristics	Reference or source
Strain		
IC762	W3110 F⁻ *his trpE3B248A2 rpsL*	19
IC1107	W3110 F⁻ *trpE3B248A2 recA rpsL*	28
JM83	ara Δ(*lac-proAB*) *rpsL* ϕ80 *lacZ*ΔM15	41
P678-54	*thr leu*	1
JM101	*supE thi* Δ(*lac-proAB*) (F′ *traD36 proAB lacI*^{q} *lacZ*ΔM15)	41
JM110	*rpsL thr leu thi lacY galK galT ara tonA tsx dam dcm supE44* Δ(*lac-proAB*) (F′ *traD36 proAB lacI*^{q} *lacZ*ΔM15)	41
Plasmid		
pBR322	Amp^r Tc^r	4
pACYC184	Tc^r Cm^r	7
pUC19	Amp^r	41
pUC18	Amp^r	41
pZAZ167	Amp^r TrpI$^+$B$^+$A$^+$	27
pZAZ131	Amp^r TrpI$^-$B$^+$A$^+$; from pZAZ167 by deletion of 368-bp *Sac*II fragment	28
pZAZ133	Amp^r TrpI$^-$B$^+$A$^+$; from pZAZ167 by deletion of 424-bp *Sac*II fragment	28
pZAZ202-489	Cm^r TrpI$^+$	This work
pZAZ202-102	Cm^r TrpI$^+$	This work
pZAZ125	Cm^r TrpI$^+$B$^+$A$^+$	28
pMI10	Cm^r TrpI$^-$B$^+$A$^+$; from pZAZ125 by deletion of *Ava*I fragment	This work
pMI40	Amp^r TrpI$^+$; *Pst*I-*Sal*I fragment subcloned into pUC19	This work
pMI50	Amp^r TrpI$^-$; *Pst*I-*Sal*I fragment subcloned into pUC18	This work
pMI110	Amp^r TrpI$^+$; from pZAZ167 by deletion of *Sph*I fragments	This work
pMIB85	Cm^r TrpI$^+$; *trpB* region removed from pZAZ202-489	This work
pMIK13	Amp^r; *Bam*HI-*Kpn*I fragment from pMI85 subcloned into pUC19	This work
pMIE6	Amp^r TrpI$^+$; *Bam*HI-*Sal*I fragment from pMI85 subcloned into pUC19	This work
pMIT3	Tc^r TrpI$^+$	This work
pUC19TET	Tc^r	This work
pBD6	Tc^r; from pMIT3 by deletion of *Bal*I-*Bal*I and *Bal*I-*Stu*I fragments	This work

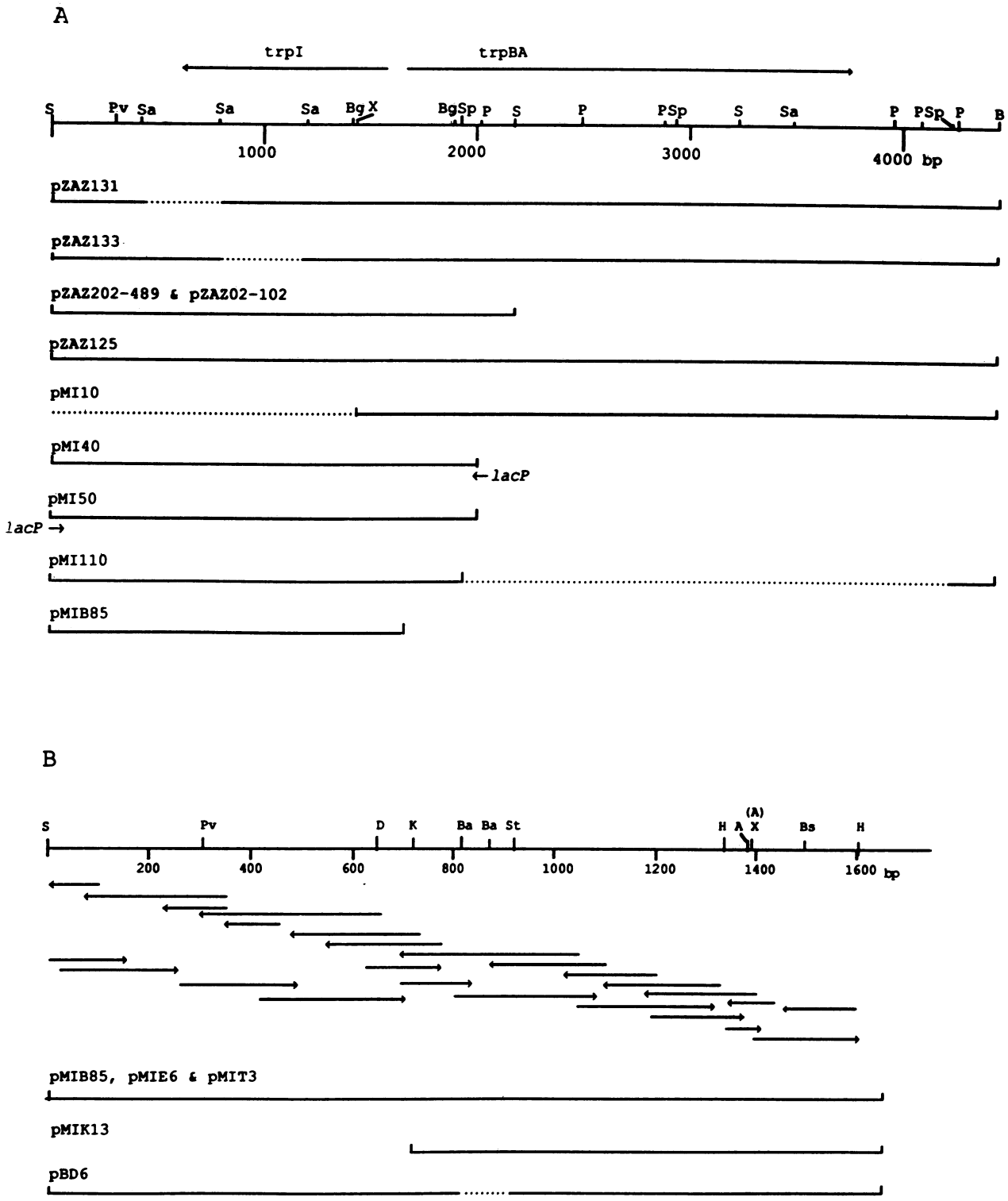


FIG. 2. (A) Restriction endonuclease cleavage map of the *trpIBA* portion of pZAZ167, with distances given in base pairs. Arrows indicate the location and orientation of the *trpBA* and *trpI* genes (18, 27, 28; this work). (B) Sequencing strategy in the *trpI* region. Nucleotides are numbered as in panel A. Below the restriction map are indicated the fragments used to obtain the sequence. Restriction enzyme abbreviations: A, *Ava*I; B, *Bam*HI; Ba, *Bal*I; Bg, *Bgl*II; Bs, *Bss*hII; D, *Dde*I; H, *Hin*fI; K, *Kpn*I; P, *Pst*I; Pv, *Pvu*I; S, *Sal*I; Sa, *Sac*II; Sp, *Sph*I; St, *Stu*I; X, *Xho*I. ——, DNA fragments retained in the corresponding plasmids; ····, regions deleted in the plasmids indicated. Relative positions of the *lacP* promoter and the *Pst*I-*Sal*I DNA fragment are shown for plasmids pMI40 and pMI50.

developed by 8% urea-polyacrylamide gel electrophoresis according to Sanger and Coulson (31). The sequence data were analyzed with the aid of the PCS computer program (23).

Minicell preparation and labeling. Plasmid-specified proteins were synthesized and labeled in minicells (8), using the *E. coli* P678-54 minicell-producing strain (1). Several steps in the usual procedure were modified as described in Results.

SDS-polyacrylamide gel electrophoresis. Labeled polypeptides were separated and their molecular weights were estimated on sodium dodecyl sulfate (SDS)-polyacrylamide gels (39). The gels were prepared for autoradiography by the technique of Bonner and Laskey (5).

Analysis of protein sequence data. The FASTP program shows the alignment of amino acid sequences and gives two similarity scores, an initial score and an optimized score, which are calculated according to the PAM250 matrix (14), an amino acid replaceability matrix. The statistical significance of a score, termed the z value, can be obtained by the equation: z = (similarity score − mean of random scores)/standard deviation of random scores. The similarity scores, mean of random scores, and standard deviation of random scores are available in the FASTP and XRDF programs, both offered by Bionet. The guidelines for z values are: $z > 3$, possibly significant; $z > 6$, probably significant; $z > 10$, significant (25).

RESULTS

Localization of the *trpI* gene by BAL 31 deletions. To study the *trpI* gene, the 2.2-kb *Sal*I-*Sal*I fragment from pZAZ167 (Fig. 2A) was subcloned into plasmid pACYC184 (7), a vector compatible with pZAZ131. Preliminary analysis of the *trpI* and *trpBA* genes on pZAZ167 suggested that this *Sal*I DNA fragment contains the first part of the *trpB* gene and the entire region upstream of *trpB*, where *trpI* is located. Plasmids pZAZ202-489 and pZAZ202-102 were generated having opposite orientations of the inserted fragment within the *tet* gene of pACYC184.

Growth rate tests showed that pZAZ202 plasmids of either orientation were TrpI$^+$. Either version of pZAZ202, when cotransformed into IC1107 with the *trpI*-deficient plasmid pZAZ131, enhanced the growth rate on MAAHC plates to the level of cells containing pZAZ167. This finding indicates that the *trpI* gene is within the 2.2-kb *Sal*I-*Sal*I DNA fragment and can be expressed from its own promoter.

To localize the 5′ and 3′ boundaries of the *trpI* gene, BAL 31 exonuclease was used to create a series of deletions invading from either end. pZAZ202-489, having a unique *Pvu*I site at the 3′ end and a unique *Pst*I site at the 5′ end (Fig. 2), was opened at the *Pvu*I site and then treated with BAL 31 exonuclease for varying times. Plasmids of various sizes were self-ligated by using T4 DNA ligase and transformed into *E. coli* IC1107. Deletion plasmids were analyzed by restriction enzyme digestion and transformed into pZAZ131-containing competent cells for growth rate tests. Results of growth rate tests suggested that TrpI activity was retained until the deletion reached the *Dde*I site at base pair (bp) 649. Five of these plasmids were chosen for sequencing because restriction enzyme mapping showed that three of them contained the largest TrpI$^+$ deletions and two had the smallest TrpI$^-$ deletions. DNA sequencing precisely defined the missing regions (Fig. 3). The largest deletion with a TrpI$^+$ phenotype, that in plasmid 58, ended at bp 615, whereas the smallest TrpI$^-$ deletion, that in plasmid 62, stopped at bp 690. Apparently one boundary of the *trpI* gene lies within this 76-bp segment, from bp 615 to 690.

The same method was used to construct a series of deletions at the opposite end. In this case, pZAZ202-489 was digested with *Pst*I to generate linear molecules for BAL 31 deletion. The results from growth rate tests and restriction enzyme mapping showed that this boundary was located between the *Hin*fI and the *Bss*hII sites. Of six deletion plasmids lacking the *Hin*fI site at bp 1601 and retaining a *Bss*hII site at bp 1494, three were phenotypically TrpI$^+$, one was intermediate between plus and minus, and two were TrpI$^-$. The ends of the deletions were clarified by DNA sequencing (Fig. 3). Except for deletion P98, these deletions are consistent with a boundary lying between bp 1544 and 1586.

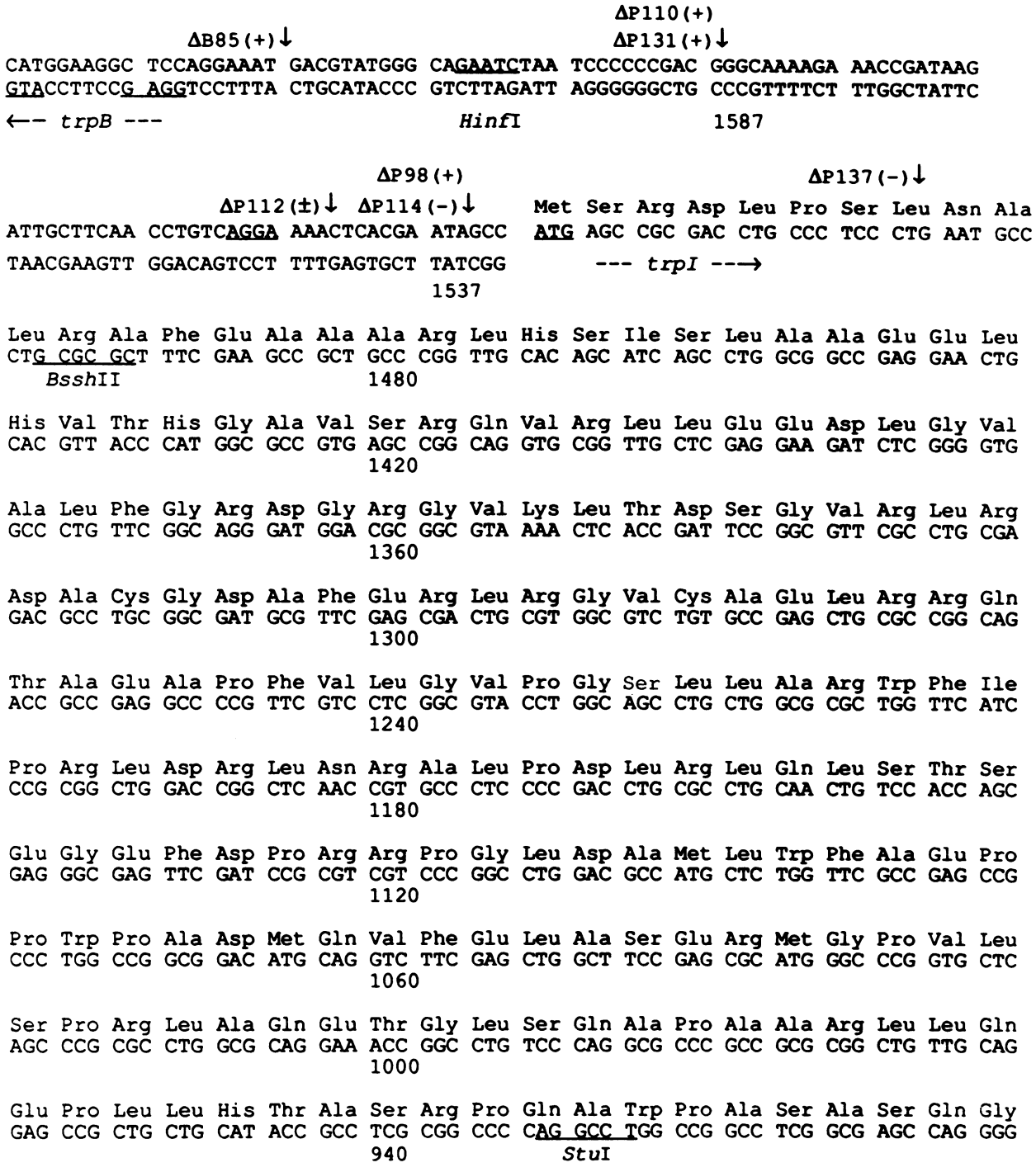
Determination of the orientation of the *trpI* gene. A 2.0-kb *Sal*I-*Pst*I fragment slightly shorter than the *Sal*I-*Sal*I fragment was inserted into the multiple cloning sites of pUC19 and pUC18 (41), generating pMI40 and pMI50, respectively. In one case, the *trpI* gene should be codirectional with the strong promoter *lacP*.

Plasmids pMI40 and pMI50 are incompatible with pZAZ131; therefore, pMI10 was made by deleting most of the *trpI* gene from pZAZ125, a pACYC184 derivative containing the entire *trpIBA* segment. The growth rate of pMI10-transformed cells on MAAHC plates was similar to that of pZAZ131-transformed cells and was increased by cotransformation with plasmid pMI110, made by deletion of the *trpBA* genes from pZAZ167.

The carbon sources used in this experiment were slightly modified from those used in previous studies (28). Three different carbon sources were added to MAHC in an attempt to increase the rate of transcription from the *lac*P promoter: 0.5% (wt/vol) sodium succinate, 0.2% (wt/vol) lactose, and 0.2% (wt/vol) glucose plus 10^{-3} M IPTG (isopropyl-β-D-thiogalactopyranoside). All three gave the same result; IC1107 containing pMI10 and pMI40 grew faster on anthranilate than did IC1107 containing pMI10 and pMI50, which indicated that in pMI40 the *trpI* gene is oriented in the same direction as the *lac*P promoter. This finding shows that the orientation of the *trpI* gene is opposite to that of *trpBA* genes. The result of one growth rate test using sodium succinate as the carbon source is presented in Table 2. It is not clear why the IC1107 cells containing pMI40 and pZAZ131 grew more slowly on MAAHC plates than did the same cells containing pZAZ167. Plasmid pMIE6 has a structure similar to that of pMI40 except that the *trpB* portion is deleted in pMIE6. When cotransformed with pZAZ131, pMIE6 was able to enhance the growth rate of IC1107 cells on MAAHC plates to the level of cells containing pZAZ167.

Determining the DNA sequence and reading frame of the *trpI* gene. A restriction map of the *trpI* region, along with the strategy of DNA sequencing, is shown in Fig. 2B. The DNA sequence of both strands was determined by the Maxam-Gilbert method (26). There is an open reading frame consistent with the *trpI* termini determined by BAL 31 deletions. The nucleotide sequence of *P. aeruginosa* PAC174 *trpI* and the deduced amino acid sequence are shown in Fig. 3.

As discussed below, there are several alternative open reading frames in the *trpI* region of pZAZ167. Making use of the gene fusion method, we identified the correct one by subcloning the *Bam*HI-*Kpn*I fragment from pMIB85, described in the next paragraph, into the multiple cloning site of pUC19 to make pMIK13. The cloning strategy is shown in Fig. 4. The DNA sequence of pMIK13 shows that translation from the *lacZ* start codon ends 2 bp ahead of the start codon of the presumed *trpI* reading frame, and after 272 codons the presumed reading frame is fused in frame with the *lacZ*



α-peptide sequence. If this presumed reading frame is expressed directly or by translational coupling, the hybrid gene should produce a fusion protein having some α-peptide activity. If another open reading frame is actually used for *trpI*, no α-peptide activity will be seen. A ligated mixture of *Bam*HI-*Kpn*I fragments and linearized pUC19 molecules was transformed into JM83. The transformed JM83 cells were grown on L agar plus X-Gal and ampicillin. Ten white and ten blue colonies were picked at random. Plasmids were analyzed by restriction enzyme digestion. Only seven plasmids, all purified from blue colonies, contained the *Bam*HI-*Kpn*I insert. This result showed that JM83 transformed with pMIK13 exhibits a blue color reaction with X-Gal, diagnostic of α-peptide activity and confirming the reading frame of the *trpI* gene.

Determination of the size of the *trpI* gene product. Because of the small amount of the *trpI* gene product in pZAZ202-489-transformed cells, it was not feasible to attempt to determine the size of the gene product directly. Plasmid pUC19, existing in high copy in the cell and containing the

```
Leu Ala Ala Glu Ala Leu Arg Tyr Gly Gln Gly Phe Glu His Leu Tyr Tyr Leu Leu Glu
CTG GCG GCG GAG GCG CTG CGC TAT GGC CAG GGG TTC GAG CAT CTC TAC TAC CTG CTG GAA
                           880 BalI

Ala Ala Val Ala Gly Leu Gly Val Ala Ile Ala Pro Glu Pro Leu Val Arg Asp Asp Leu
GCG GCG GTG GCG GGC CTC GGC GTG GCC ATC GCC CCG GAG CCA CTG GTC CGC GAC GAT CTC
                           820 BalI

Ala Ala Ala Pro Gly Gly Pro Trp Gly Phe Ile Glu Thr Asp Ala Arg Leu Ala Leu Trp
GCC GCG GCG CCT GGC GGG CCC TGG GGC TTC ATC GAG ACC GAC GCG CGC CTG GCC CTG TGG
                           760
                          ↓Δ47(-)      ↓Δ62(-)
Val Pro Ala Arg Leu His Asp Pro Arg Ala Gly Arg Leu Ala Gln Trp Leu Arg Glu Gln
GTA CCG GCG CGC CTG CAC GAT CCG CGT GCC GGG CGC CTG GCG CAA TGG TTG CGG GAG CAG
 KpnI                     700

Leu Ala Gly  *
TTG GCA GGC TGA
          DdeI                                                    Δ46(+)
                                        ↓Δ58(+)                  ↓Δ52(+)
 GCGCTGGCG ATTTGCTGCA CTTTTCCGGC TCGGCTAGTG TCCATTGAAC AGAGGCGCCC GTCGGGCGCG
 649                                                    600

AAGAACCAGG CATCGTCCGG GGAGAGAGTT CATGTCCAAT CACCACACCT ACAAGAAGAT CGAACTGGTC
                           550

GGCTCGTCCA AGACCAGCAT CGAGGACGCC ATCAACAACG CCCTCGCCGA AGCGGCGAAG AGCATCCAGC
          500                                                    450

ATCTGGAATG GTTCGAGGTG GTGGATACCC GCGGGCACAT CGAGAACGGT GCCGTCGGCC ATTACCAGGT
                                    400

GACCCTGAAA GTAGGGTTCC GTATCGCCAA TAGCTGATGG CACCCTGGAA ACCTGTCCAC GATCGCCTGC
                     350                                          PvuI

GCCCTGGAAA GGCCGCCGCG TCTTTCCGCT TCGCCGGATC ATGCACAGGC TATGAGGGCG GCGTGCCTCC
300                                                    250

ATGGGCGGCG CGCTCGTCTC AACCTTGCCG AATTTGGCAA GCCTGACCCG TTCGGGTCTT CATCCATGCA
                           200

ACGAGGGAAT GTGTTCATGA AGAAGTTGAT GTTGGCAGTC GGCCTGTTTG CCGTGGCGGG CAGCGCATTC
          150                                                    100

GCCGCCAAGC CCTGTGAGGA ACTGAAAGCC GAGATCGATG CGAAGATCAA GGCCAACGGC GTTCCTGCCT
                                    50

ACACCCTGGA AATCGTCGAC
                     1
                   SalI
```

FIG. 3. Nucleotide sequence of the *trpI* gene from *P. aeruginosa*, 5′ and 3′ flanking regions of the gene, and amino acid sequence of the *trpI* gene product. The *trpB* and *trpI* genes are transcribed divergently (dashed arrows). Initiation codons and putative ribosome-binding sequences complementary to 16S rRNA are underlined. The BAL 31-deleted plasmids and their deletion endpoints are indicated above the sequence. For the 5′-end deletions, nucleotides to the right of the endpoint are retrained; for the 3′-end deletion, nucleotides to the left of the endpoint are retained. (+), (−), Designations of the phenotype of each deletion. Several restriction enzyme sites mentioned in the text are indicated below the sequence.

strong *lacP* promoter, was used to elevate the amount of the *trpI* gene product in the cell. Since no convenient restriction enzyme sites were available to subclone the *trpI* gene into pUC19, the BAL 31 enzyme was used to delete the *trpB* region in pZAZ202-489. The ends of the BAL 31-digested molecules were made blunt by DNA polymerase I large fragment in the presence of all four deoxynucleotide triphosphates and then ligated to *Bam*HI linkers (CGGATCCG). Plasmid pMIB85 having a *Bam*HI restriction enzyme site located 87 nucleotides ahead of the presumed *trpI* start codon, was constructed in this way. The *Bam*HI-*Sal*I DNA fragment from pMIB85 was cloned into pUC19 to generate pMIE6.

Plasmid-directed polypeptides synthesized in *E. coli* minicells and separated on SDS-polyacrylamide gels were used to examine the size of the *trpI* protein. The preliminary

TABLE 2. Growth rate test to determine the orientation of the *trpI* gene

Cell (plasmid)	Colony size on[a]:							
	MTAHC		MAAHC		MIAHC		MAHC	
	22 h	34 h	22 h	34 h	22 h	34 h	22 h	34 h
IC1107(pZAZ167)	+	++	+	++	±	±	−	−
IC1107(pZAZ125)	+	++	+	++	−	−	−	−
IC1107(pMI10)	+	++	−	±	−	−	−	−
IC1107(pMI10, pMI110)	+	++	+	++	±	±	−	−
IC1107(pMI10, pMI40)	+	++	+*	+	±	±	−	−
IC1107(pMI10, pMI50)	+	++	−	−	−	−	−	−
IC1107(pMI10, pMIE6)	+	++	+	++	±	±	−	−

[a] Media are described in Materials and Methods. Sodium succinate was used as the carbon source. The first reading was 22 h after innoculation, and the second was 12 h later. Colony size: −, single colonies barely visible to the naked eye; ±, colonies about 0.1 to 0.2 mm in diameter, easily visible to the naked eye; +*, colonies about 0.5 mm in diameter; +, colonies about 1 mm in diameter; ++, colonies about 1.5 mm in diameter.

experiment showed that among the pMIE6-directed polypeptides, the *trpI* and β-lactamase polypeptides comigrated on SDS-polyacrylamide gels. The size of the *trpI* gene product predicted from translation of its DNA sequence is 31,950 daltons. The β-lactamase encoded by the ampicillin resistance gene in the pUC19 derivative has two forms in minicells; the sizes of the unprocessed and processed β-lactamases are 31,410 and 28,850 daltons, respectively (2, 36). To visualize the *trpI* polypeptide, the ampicillin resistance gene was removed and replaced by the tetracycline resistance gene for selection. A portion of the ampicillin resistance gene was deleted, and a tetracycline gene cartridge, the *Eco*RI-*Ava*I DNA fragment from pBR322, was inserted in its place, resulting in pMIT3. Plasmid pMIT3 was transformed into the minicell-producing strain *E. coli* K-12 P678-54 for in vivo translation. Suitable amounts of minicell extracts were loaded on a 10% SDS-polyacrylamide gel (Fig. 5). Comparison of the polypeptides produced from pMIT3 and pUC19TET shows that there is one extra band having a molecular weight of about 29,000 and migrating to the position of the two β-lactamase bands in pUC19. When produced from the deletion mutant pBD6, this polypeptide is

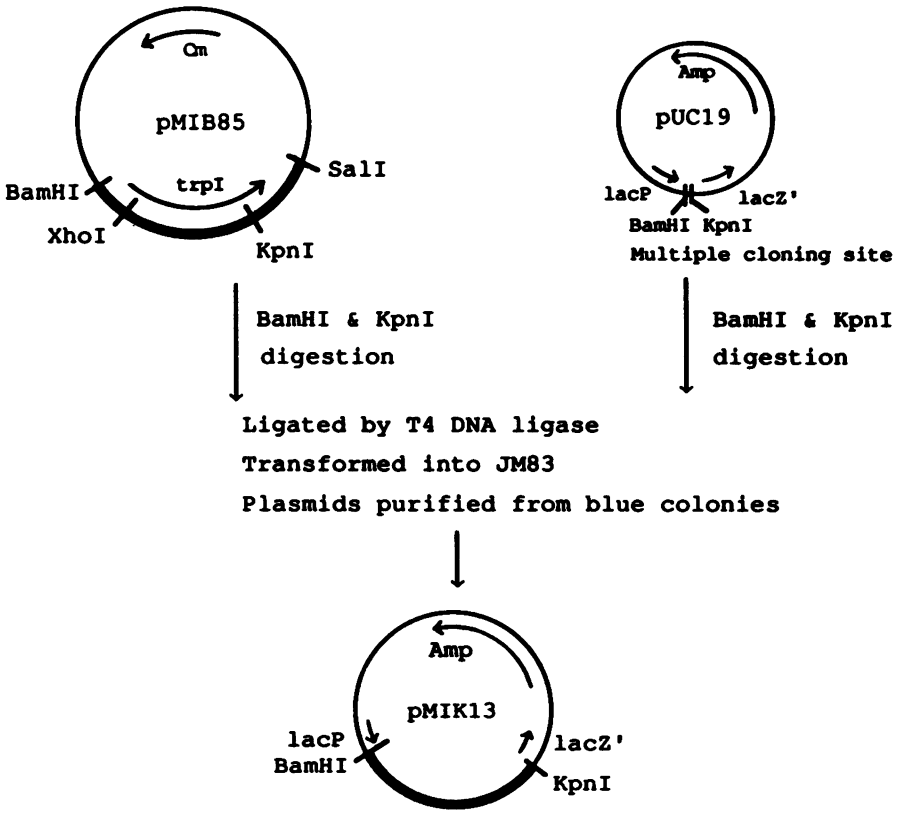


FIG. 4. Construction of plasmid pMIK13. The *Bam*HI-*Kpn*I fragment, containing nearly 90% of the *trpI* gene, is positioned downstream of the *lacP* promoter so that the orientation of the *trpI* gene is the same as the direction of transcription from the *lacP* promoter. The multiple cloning site of pUC19, a 54-bp synthetic oligonucleotide, is inserted between the sixth and seventh codons of *lacZ'*. *lacZ'* encodes the N-terminal 145 residues of β-galactosidase, termed the α-peptide, whose activity is not affected by insertion of the multiple cloning site. The *trpI* gene region is indicated by heavy lines. Construction of pMIB85 is described in Results.

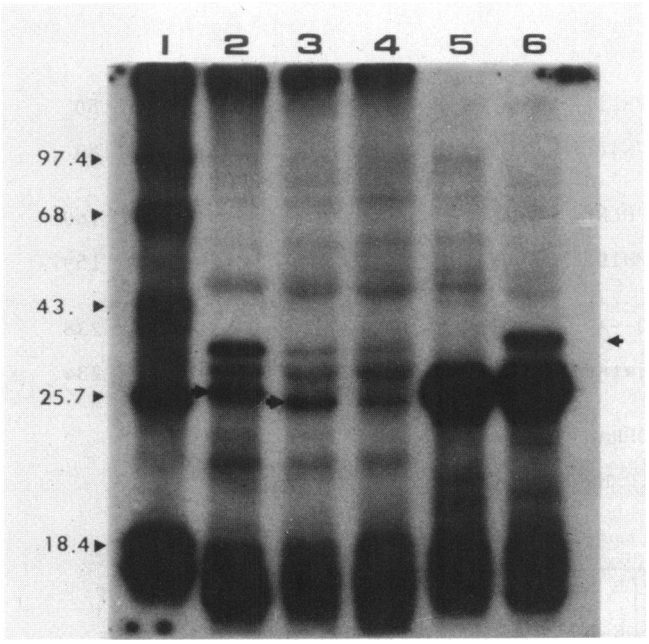


FIG. 5. Polypeptides produced by pMIT3- and pBD6-containing minicells. Polypeptides were labeled in minicells by [^{35}S]methionine plus ^{14}C-amino acids and separated on a 10% polyacrylamide gel in the presence of SDS. Lane 1, ^{14}C-labeled protein markers (97.4, 68, 43, 25.7, and 18.4 kilodaltons). Polypeptides specified by the following plasmids are shown: lane 2, pMIT3; lane 3, pBD6; lane 4, pUC19tet; lane 5, pUC19; lane 6, pMIK13. Plasmid pUC19tet, a pUC19 derivative, was constructed by the method used to construct plasmid pMIT3. The arrows beside lanes 2, 3, and 6 indicate the polypeptides encoded by the intact *trpI* gene, the defective *trpI* gene, and the *trpI-lacZ'* hybrid gene, respectively.

proportionally decreased in size (the construction of pBD6 is described in Materials and Methods). Apparently, this extra band is the product of the *trpI* gene.

Similarity among the *trpI*, *ampR*, *ilvY*, *metR*, *lysR*, and *nodD* proteins. With the help of the FASTP program (25), the *trpI* amino acid sequence was compared with protein sequences in the PIR (Protein Identification Resource) and SWISS-PRO (Swiss Protein) data bases. *ampR* in *Enterobacter cloacae* showed a significant similarity through the entire amino acid sequence (Fig. 6A). *lysR* in *E. coli* (35) and *nodD* in *R. meliloti* (16) were found to have some similarity to the *trpI* protein in their N-terminal portions. Subsequently, the *metR* (30) and *ilvY* (40) proteins in *E. coli* were examined and also showed similarity in this region. These six proteins, all functioning as transcriptional activators, had various degrees of similarity along their amino acid sequences. It is interesting that they are all related in the N-terminal regions (Fig. 6B). The order of their listing in the figure indicates their relatedness based on similarity scores calculated by the XRDF program (data not shown). No gaps were introduced into this region for alignment. If gaps are allowed, several additional small regions can be matched in other parts of the proteins. However, not all of these small regions are common to all six proteins. Analysis of similarity scores indicates that the N-terminal portions of each of the adjacent pairs are significantly similar (based on the initial z value; see Materials and Methods).

DISCUSSION

The approximate location of the *trpI* gene was established by genetic studies using low-level, constitutive mutants of pZAZ167; two such derivatives, pZAZ131 and pZAZ133, are deleted in different but adjacent *Sac*II fragments. This finding suggests that the *trpI* gene may extend over these two *Sac*II fragments (28). When sequencing in the vicinity of *trpI*, upstream of the *trpB* gene in pZAZ167, we found this region to have a very high G+C content, resulting in severe compression in several areas of both strands of the DNA sequence. (The average G+C content in the genome of *P. aeruginosa* is reported to be 65% [32], and that in the *trpBA* gene pair is 68% [18]). The 72% G+C content in the *trpI* gene region also results in a low frequency of occurrence of stop codons and a corresponding abundance of open reading frames on both strands of the DNA sequence. Consequently, we found it necessary to use deletions generated by BAL 31 nuclease to better determine the boundaries of the *trpI* gene.

The results of BAL 31 deletion experiments showed that the 3' boundary of the *trpI* gene lies between bp 615 and 690; the 5' boundary is located between bp 1544 and 1586. Thus, the *trpI* gene extends from about 60 bp upstream of the *trpB* start codon to the middle of one *Sac*II fragment (Fig. 3), consistent with the findings of Manch and Crawford (28). P98 and P114, 5'-end deletion mutants, have identical sequences in the *trpI* gene region but differ in the extent of deletion into the adjacent *tet* gene. The difference in phenotype of these two deletions may be caused by the upstream flanking sequence. The orientation of the *trpI* gene in pZAZ202-489 is opposite to the transcriptional direction of the *tet* gene; therefore, it is unlikely that expression of the *trpI* gene is affected by the promoter of the *tet* gene. Close and Rodriguez (9) used a chloramphenicol resistance gene cartridge to map the transcriptional activities of pBR322. They found that a low level of transcription occurs in the antisense direction within the *tet* gene. This transcriptional activity may extend to the *trpI* gene in deletion P98.

The orientation of *trpI* is opposite to that of the *trpBA* gene pair. Such an arrangement, in which a regulatory gene is transcribed divergently from the gene set that it regulates, has been observed before (24). These two genes or operons can share a common control region. In some cases, the two start sites are so close that the promoters overlap, as with *araC-araBAD* (24) and *ilvY-ilvC* in *E. coli* (40) and *nahR-nahG* in the *Pseudomonas* NAH7 plasmid (33). In the *trpI* and *trpBA* system, the 103-bp distance between the two initiation codons implies that the promoters may partially overlap (Fig. 3).

Growth rate tests were used to examine the activity of the *trpI* gene product in transformed cells. These tests do not, however, precisely measure the amount of the *trpI* gene product. Once enough endogenous tryptophan is made to supply the cell, the level of tryptophan synthase, the product of the *trpBA* gene pair, is not limiting for growth. Plasmid pMIK13 was used to confirm the reading frame of *trpI* predicted from the DNA sequence. The *trpI-lacZ'* fusion protein generated in pMIK13-containing cells can also be used to monitor the level of the *trpI* gene product. We found that transformed cells cultured in glycerol plus IPTG had the highest level of β-galactosidase activity. This culture condition was applied to the minicell system to generate enough of the *trpI* protein to be seen on SDS-polyacrylamide gels.

One open reading frame, encoding 293 residues found within the *trpI* boundaries, was confirmed to be expressed in vivo. In recombinant plasmid pMIK13, this reading frame was fused in frame with the α-peptide sequence. The α-peptide activity was detected in pMIK13-transformed cells, which indicated that this fusion protein was synthesized in vivo (Fig. 5). In addition, a full-sized fusion polypeptide was

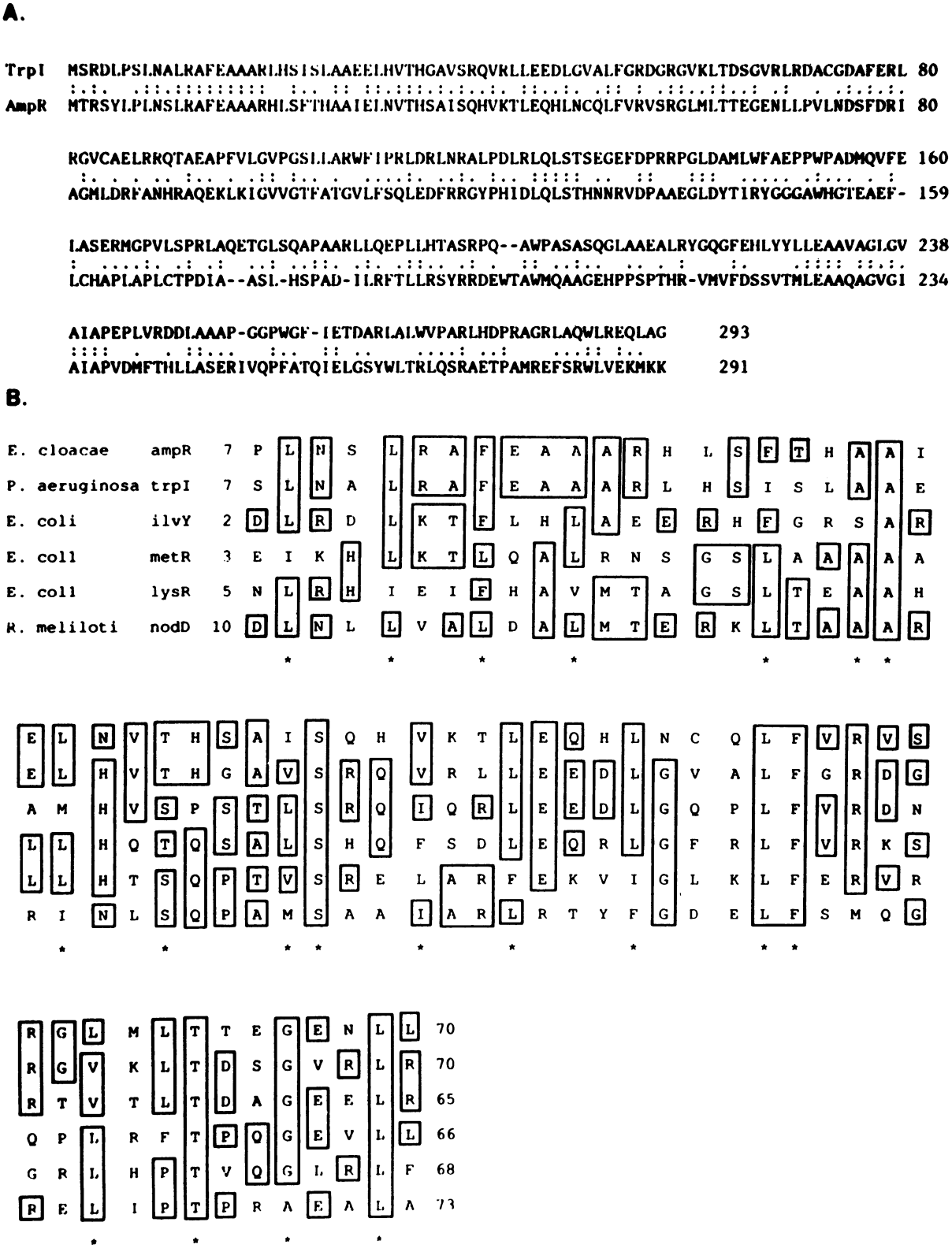


FIG. 6. (A) Alignment of the *trpI* and *ampR* amino acid sequences, performed by the FASTP program. Symbols: :, identical residues; ., residues frequently substituted in evolution; -, insertions made during optimization. The similarity between the *trpI* and *ampR* proteins was evaluated by the *z* value, calculated by the XRDF program. The optimized similarity score of the *trpI-ampR* pair is 374. The mean of random scores and standard deviation of random scores were 38.3 and 8.59, respectively. Therefore, the *z* value is significant at 39.08 (see Materials and Methods). (B) Alignment of the N-terminal portions of the *trpI*, *ampR*, *ilvY*, *metR*, *lysR*, and *nodD* amino acid sequences. Residue numbers are given at the beginning and end of each sequence. Positions with identical residues are boxed; those with residues similar or identical are starred. Similarity groups: D = E; F = Y; N = Q; K = R; S = T; and I = L = V = F = M.

TABLE 3. Codon usage in *P. aeruginosa trpI*

Amino acid	Codon	Usage	Amino acid	Codon	Usage	Amino acid	Codon	Usage	Amino acid	Codon	Usage
Phe	UUU	0	Ser	UCU	0	Tyr	UAU	1	Cys	UGU	1
Phe	UUC	10	Ser	UCC	5	Tyr	UAC	2	Cys	UGC	1
Leu	UUA	0	Ser	UCA	0	END	UAA	0	END	UGA	1
Leu	UUG	5	Ser	UCG	2	END	UAG	0	Trp	UGG	7
Leu	CUU	0	Pro	CCU	2	His	CAU	3	Arg	CGU	5
Leu	CUC	11	Pro	CCC	7	His	CAC	3	Arg	CGC	14
Leu	CUA	0	Pro	CCA	1	Gln	CAA	2	Arg	CGA	2
Leu	CUG	31	Pro	CCG	12	Gln	CAG	10	Arg	CGG	9
Ile	AUU	0	Thr	ACU	0	Asn	AAU	1	Ser	AGU	0
Ile	AUC	4	Thr	ACC	7	Asn	AAC	1	Ser	AGC	8
Ile	AUA	0	Thr	ACA	0	Lys	AAA	1	Arg	AGA	0
Met	AUG	4	Thr	ACG	0	Lys	AAG	0	Arg	AGG	1
Val	GUU	2	Ala	GCU	3	Asp	GAU	7	Gly	GGU	0
Val	GUC	4	Ala	GCC	22	Asp	GAC	8	Gly	GGC	18
Val	GUA	3	Ala	GCA	1	Glu	GAA	5	Gly	GGA	1
Val	GUG	6	Ala	GCG	19	Glu	GAG	16	Gly	GGG	5

visible on SDS-polyacrylamide gels after synthesis in the minicell system. The sizes of both the fusion polypeptide and the intact *trpI* polypeptide estimated from SDS-polyacrylamide gels were close to those predicted from the amino acid sequence. Deletion of 37 residues in the middle of this reading frame resulted in loss of the activating function of the *trpI* protein (data not shown) and decreased the size of the *trpI* polypeptide in SDS-polyacrylamide gels.

The codon usage in *trpI*, similar to that of other *Pseudomonas* genes, shows a marked preference for G or C in the third position (Table 3) (12, 18). As noticed when *P. aeruginosa trpBA* was compared with *trpBA* from other bacteria, lysine is used less frequently, usually being substituted for by arginine (18); in the *trpI* gene arginine is also much preferred, composing 10.6% of the total residues, whereas histidine and lysine constitute 2.0 and 0.3%, respectively. In addition, the preference for leucine over isoleucine is remarkable in comparison with the case in other *Pseudomonas* genes; leucine accounts for 16% and isoleucine accounts for only 1.4% of the total residues. The amino acid composition of *trpI* is shown in Table 4. The slightly higher G+C content (72% versus about 65%) and biased usage of certain amino acids in the *trpI* gene might be accounted for by the feeble expression of this gene, as is true of regulatory genes in general.

The *trpI*, *ampR*, *ilvY*, *metR*, *lysR*, and *nodD* gene products shown in Fig. 6B have some common features. They function as activators, required for copious expression of other genes. In the presence of the β-lactam antibiotic cefoxitin, the *ampR* product induces the expression of *ampC*, encoding cephalosporinase. The *ilvY* protein activates the expression of *ilvC*, encoding acetohydroxy acid isomeroreductase, the second enzyme in the isoleucine-valine biosynthetic pathway (40). The *metR* product is required for expression of the *metE* and *metH* gene products, enzymes responsible for the methylation of homocysteine to methionine (30). The *lysR* protein activates the *lysA* gene, encoding diaminopimelate decarboxylase (35). In the presence of root extract, the *nodD* protein induces the *nodABC* genes, which are involved in root invasion and the stimulation of nodule development (16, 17). In some cases these activators may serve to repress their own synthesis. All of these activators are similar in size: the *trpI*, *ampR*, *ilvY*, *lysR*, and *nodD* proteins have 293, 291, 297, 311, and 308 residues, respectively, whereas the *metR* protein has 276 residues. These regulatory genes are all transcribed divergently from the gene sets that they regulate, and they often lie close enough that the divergent promoters are overlapping. For the *lysR-lysA* pair, the start codons are separated by 121 bp; the start codon of *nodD* is 266 bp from that of *nodA*. For the *ampR-ampC* pair, the transcription initiation sites are separated by 55 bp; those of the *metR-metE* pair and the *ilvY-ilvC* pair are 25 and 45 bp apart, respectively.

The significant similarity in N-terminal portions in this family of activators probably reflects a common function among them. In the case of the catabolite activator protein, the N-terminal domain contains the cyclic AMP binding site and the C-terminal domain has the DNA-binding region, which recognizes a specific DNA sequence and also interacts with RNA polymerase (22). In these six proteins the conserved N-terminal portions may be the DNA-binding regions. The diversity of their C-terminal portions may relate to inducer specificity, which suggests that this portion may contain the inducer-binding site. A 20-residue sequence within the N-terminal portion of *lysR* has been proposed to

TABLE 4. Deduced amino acid composition of the *P. aeruginosa trpI* gene product[a]

Amino acid	Composition	
	No. of residues	%
Ala	45	15.4
Arg	31	10.6
Asn	2	0.7
Asp	15	5.1
Cys	2	0.7
Gln	12	4.1
Glu	21	7.2
Gly	24	8.2
His	6	2.0
Ile	4	1.4
Leu	47	16.0
Lys	1	0.3
Met	4	1.4
Phe	10	3.4
Pro	22	7.5
Ser	15	5.1
Thr	7	2.4
Trp	7	2.4
Tyr	3	1.0
Val	15	5.1

[a] Total number of residues, 293; molecular weight, 31,950.

be a *cro*-like DNA-binding region; His-29 and its adjacent residues serve as the turn (15). The alignment of the *trpI* protein with the *lysR* protein offers another prediction for the location of a helix-turn-helix; that is, Gly-35 in the *trpI* protein and serine or proline for the other proteins may serve as the center of the turn. The *ilvY* member has been proven to be a DNA-binding protein, and RNA polymerase binding to the *ilvC* promoter is mediated by the *ilvY* protein in the presence of the acetohydroxy acid isomeroreductase substrate (40a). In all likelihood, this entire family of activators affects the expression of other sets of genes after binding to a specific region of DNA.

Independently, Appelbaum et al. (3) found the similarity between *Rhizobium japonicum nodD2* and *E. coli lysR*. The *R. japonicum nodD2* gene product is about 69% identical to that of *R. meliloti nodD*. Appelbaum et al. also compared the amino acid sequences of the *nodD1* and *nodD2* gene products from *R. japonicum* USDA 191 to those of the related *nodD* gene products from *R. meliloti*, *Rhizobium trifolii*, *Rhizobium leguminosarum*, and *Bradyrhizobium* ("*Parasponia*") sp. Interestingly, they found that the N-terminal regions had the greatest similarity; 60 of the first 80 residues were identical in all six genes. Spaink and co-workers (34) concluded that although the various *nodD* gene products are functionally interchangeable, they are not identical. Horvath and co-workers (21) propose that the *nodD* genes may encode determinants of host specificity by interacting with different plant factors. Comparison of amino acid sequences of the *nodD* family shows a similar pattern in that they are closely related in the N-terminal portions and more diverse in the C-terminal portions. This agrees with our proposal for the structural plan for all of the *trpI*-related activators; the N-terminal portion is responsible for DNA binding, and the C-terminal portion interacts with the inducer.

Recently, the N-terminal portion of the *catM* repressor of *Acinetobacter calcoaceticus* was shown to share apparent similarity with the six proteins mentioned above (E. L. Neidle, personal communication). This 251-residue regulatory protein controls the induction of a series of enzymes for catechol degradation; its transcription diverges from one of the operons that it regulates. Genetic tests show clearly, however, that *catM* encodes a repressor, not an activator. This finding further suggests that this family of regulatory proteins may have evolved to express varied functions, while its N-terminal portion still retains specific DNA-binding ability.

The *trpI*-related family of regulatory proteins has recently been independently described by Henikoff et al. (20). Using GENPRO software to search the EMBL, GENBANK, and NBR-PIF data bases, they identified three additional examples from *E. coli* and *Salmonella typhimurium* and one from *Alcaligenes eutrophus*. Those from the enteric bacteria included *cysB*, an activator in the cysteine synthetic pathway, and two genes of unknown function, *leuO*, located between the *leuACBD* and *ilvIH* operons, and *antO*, located upstream of the gene for the Na^+/H^+ antiporter. The *Alcaligenes* example is also an open reading frame of unknown function found upstream of the *tfdA* (2,4-dichlorophenoxyacetate monooxygenase) gene. Henikoff et al. (20) proposed a helix-turn-helix motif located in the N-terminal region in each of the family members they identified. They found no sequence similarity between these regulatory proteins and those of the *ompR* family.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI20279 from the National Institutes of Health. Computer resources for the data base search were provided by the National Institutes of Health-sponsored BIONET National Computer Resource for Molecular Biology (Public Health Service grant 1 U41 RRO1685).

We thank R. Milkman for critically reading the manuscript and E. L. Neidle and L. N. Ornston for communicating their results to us prior to publication.

ADDENDUM IN PROOF

These data will appear in the EMBL/GenBank/DDBJ nucleotide sequence data bases under accession number M21093.

LITERATURE CITED

1. **Adler, H. I., W. D. Fisher, A. Cohen, and A. Hardigree.** 1967. Miniature *Escherichia coli* cells deficient in DNA. Proc. Natl. Acad. Sci. USA **57**:321–326.
2. **Ambler, R. P., and G. K. Scott.** 1978. Partial amino acid sequence of the ampicillin resistance gene of *Escherichia coli* plasmid R6K. Proc. Natl. Acad. Sci. USA **75**:3732–3736.
3. **Appelbaum, E. R., D. V. Thompson, K. Idler, and N. Chartrain.** 1988. *Rhizobium japonicum* USDA191 has two *nodD* genes that differ in primary structure and function. J. Bacteriol. **170**:12–20.
4. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Bettach, H. L. Heynecker, H. W. Boyer, J. H. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
5. **Bonner, W. H., and R. A. Laskey.** 1974. A film detection method for tritium-labeled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. **46**:83–88.
6. **Calhoun, D. H., D. L. Pierson, and R. A. Jensen.** 1973. The regulation of tryptophan biosynthesis in *Pseudomonas aeruginosa*. Mol. Gen. Genet. **121**:117–132.
7. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134**: 1141–1156.
8. **Clark-Curtiss, J. F., and R. Curtiss III.** 1983. Analysis of recombinant DNA using *Escherichia coli* minicells. Methods Enzymol. **101**:347–362.
9. **Close, T. J., and R. L. Rodriguez.** 1982. Construction and characterization of chloramphenicol-resistance gene cartridge: a new approach to the transcription mapping of extrachromosomal elements. Gene **20**:305–316.
10. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1973. Construction of biologically functional bacterial plasmids *in vitro*. Proc. Natl. Acad. Sci. USA **70**:3240–3244.
11. **Crawford, I. P.** 1987. The synthesis of tryptophan from chorismate: comparative aspects. Methods Enzymol. **142**:293–300.
12. **Crawford, I. P., and L. Eberly.** 1986. Structure and regulation of the anthranilate synthase genes in *Pseudomonas aeruginosa*. I. Sequence of *trpG* encoding the glutamine amidotransferase subunit. Mol. Biol. Evol. **3**:436–448.
13. **Crawford, I. P., and I. C. Gunsalus.** 1966. Inducibility of tryptophan synthase in *Pseudomonas putida*. Proc. Natl. Acad. Sci. USA **56**:717–724.
14. **Dayhoff, M.** 1978. Atlas of protein sequence and structure, vol. 5, suppl. 3. National Biomedical Research Foundation, Silver Spring, Md.
15. **Dodd, I. B., and J. B. Egan.** 1987. Systematic method for the detection of potential λ Cro-like DNA-binding regions in proteins. J. Mol. Biol. **194**:557–564.
16. **Eglehoff, T. T., R. F. Fisher, T. W. Jacobs, J. T. Mulligan, and S. R. Long.** 1985. Nucleotide sequence of *Rhizobium meliloti* 1021 nodulation genes: *nodD* is read divergently from *nodABC*. DNA **4**:241–248.
17. **Fisher, R. F., H. L. Brierley, J. T. Mulligan, and S. R. Long.** 1987. Transcription of *Rhizobium meliloti* nodulation genes. J. Biol. Chem. **262**:6849–6855.

18. **Hadero, A., and I. P. Crawford.** 1986. Nucleotide sequence of the genes for tryptophan synthase in *Pseudomonas aeruginosa*. Mol. Biol. Evol. **3**:191–204.
19. **Hedges, R. W., A. E. Jacob, and I. P. Crawford.** 1977. Wide ranging plasmid bearing the *Pseudomonas aeruginosa* tryptophan synthase genes. Nature **267**:283–284.
20. **Henikoff, S., G. W. Haughn, J. M. Calvo, and J. C. Wallace.** 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA **85**:6602–6606.
21. **Horvath, B., C. Bachem, J. Schell, and A. Kondorosi.** 1987. Host-specific regulation of nodulation gene in *Rhizobium* is mediated by a plant-signal, interacting with the *nodD* gene product. EMBO J. **6**:841–848.
22. **Irwin, N., and M. Ptashne.** 1987. Mutants of the catabolite activator protein of *Escherichia coli* that are specifically deficient in the gene-activation function. Proc. Natl. Acad. Sci. USA **84**:8315–8319.
23. **Lagrimini, L. M., S. T. Brentano, and J. E. Donelson.** 1984. A DNA sequence analysis package for the IBM personal computer. Nucleic Acids Res. **12**:605–614.
24. **Lee, N. L. W. O. Gielow, and R. G. Wallace.** 1981. Mechanism of *araC* autoregulation and the domains of two overlapping promoters, P_C and P_{BAD} in the L-arabinose regulatory region of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **78**:752–756.
25. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
26. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
27. **Manch, J. N., and I. P. Crawford.** 1981. Ordering tryptophan synthase genes of *Pseudomonas aeruginosa* by cloning in *Escherichia coli* J. Bacteriol. **146**:102–107.
28. **Manch, J. N., and I. P. Crawford.** 1982. Genetic evidence for a positive regulatory factor mediating induction on the tryptophan pathway of *Pseudomonas aeruginosa*. J. Mol. Biol. **156**:67–77.
29. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Plamann, L. S., and G. V. Stauffer.** 1987. Nucleotide sequence of the *Salmonella typhimurium metR* gene and the *metR-metE* control region. J. Bacteriol. **169**:3932–3937.
31. **Sanger, F., and A. R. Coulson.** 1978. The use of thin acrylamide gels for DNA sequencing. FEBS Lett. **87**:107–110.
32. **Shapiro, H. S.** 1968. Distribution of purine and pyrimidines in deoxynucleic acids, p. H-31–H-36. *In* H. A. Sober (ed.), Handbook of biochemistry: selected data for molecular biology. CRC Press, Inc., Cleveland.
33. **Schell, M. A.** 1986. Homology between nucleotide sequences of promoter regions of *nah* and *gal* operons of NAH7 plasmid of *Pseudomonas putida*. Proc. Natl. Acad. Sci. USA **83**:369–373.
34. **Spaink, H. P., C. A. Wijffelman, I. Pees, R. J. H. Okker, and B. J. J. Lugtenberg.** 1987. *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. Nature (London) **328**:337–340.
35. **Stragier, P., and J.-C. Patte.** 1983. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli*. III. Nucleotide sequence and regulation of the *lysR* gene. J. Mol. Biol. **168**:333–350.
36. **Sutcliff, J. G.** 1978. Nucleotide sequence of the ampicillin resistance gene of *Escherichia coli* plasmid pBR322. Proc. Natl. Acad. Sci. USA **75**:3737–3741.
37. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system from insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
38. **Vogel, H. J., and D. M. Bonner.** 1956. Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. **218**:97–106.
39. **Weber, K., and M. Osborn.** 1969. The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. J. Biol. Chem. **244**:4406–4412.
40. **Wek, R. C., and G. W. Hatfield.** 1986. Nucleotide sequence and *in vivo* expression of *ilvY* and *ilvC* genes in *Escherichia coli* K12. J. Biol. Chem. **261**:2441–2450.
40a. **Wek, R. C., and G. W. Hatfield.** 1988. Transcriptional activation at adjacent operators in the divergent-overlapping *ilvY* and *ilvC* promoters of *Escherichia coli*. J. Mol. Biol. **203**:643–663.
41. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
42. **Yousaf, S. I., A. R. Carroll, and B. E. Clarke.** 1984. A new and improved method for 3′-end labeling DNA using [α-^{32}P]ddATP. Gene **27**:309–313.