JOURNAL OF BACTERIOLOGY, Jan. 1990, p. 397–410
0021-9193/90/010397-14$02.00/0

Vol. 172, No. 1

Chromosomal Organization and Expression of *Escherichia coli pabA*

PHUC V. TRAN, TODD A. BANNOR,† STELLA Z. DOKTOR, AND BRIAN P. NICHOLS*

Laboratory for Molecular Biology, Department of Biological Sciences, University of Illinois at Chicago, P.O. Box 4348, Chicago, Illinois 60680

Received 17 May 1989/Accepted 12 October 1989

The *pabA* gene in *Escherichia coli* and *Salmonella typhimurium* encodes the glutamine amidotransferase subunit of *para*-aminobenzoate synthase, which catalyzes the first reaction in the conversion of chorismate to *para*-aminobenzoate (PABA). We have determined the nucleotide sequences of 1,362 base pairs preceding *E. coli pabA* and of 981 base pairs preceding *S. typhimurium pabA*. The nucleotide sequences suggest the presence of two protein-coding regions immediately upstream of *pabA*, designated *orf1* and *fic*. Transcription analysis indicates that *E. coli pabA* is encoded by two overlapping transcriptional units. The polycistronic transcriptional unit includes *orf1-fic-pabA* and is initiated by the promoter designated P_2. The monocistronic unit includes only *pabA* and is initiated by the promoter designated P_1, which is located in the *fic*-coding region. Both promoters transcribe *pabA* to about the same steady-state level. However, expression analysis using chromosomal *pabA-lacZ* translational fusions indicated that P_1 expressed PabA at least 50-fold more efficiently than P_2. *pabA*-dependent growth rate analysis indicates that P_1 is essential and P_2 is dispensable for PABA metabolism. In the absence of P_1, growth was reduced as a result of insufficient PabA expressed from P_2. The significance of these results and possible posttranscriptional control mechanisms which affect PabA expression from the P_2-initiated polycistronic unit are discussed.

para-Aminobenzoate synthase (PABS) catalyzes the conversion of chorismate and glutamine to an unidentified intermediate which is converted to *para*-aminobenzoate (PABA) by a second enzyme (29a). The two subunits of PABS are encoded by two unlinked genes: *pabB*, at 40 min on the *Escherichia coli* map (15, 43), which encodes the larger 54,000-molecular-weight subunit that catalyzes the synthesis of the unidentified intermediate from chorismate and NH_4^+ (12, 39), and *pabA*, at 74 min (14, 15), which encodes a 21,000-molecular weight glutamine-amidotransferase subunit that allows PABS to use glutamine as the source of NH_4^+ (19). The gene(s) encoding the second enzyme has not been identified or mapped.

The regulation of the genes encoding PABS is unknown. PABA is essential for the biosynthesis of dihydrofolate, which in various forms participates in the synthesis of purines, pyrimidines, formylmethionyl-tRNA, and some amino acids and vitamins. As the folate moiety is not consumed during the reactions, the de novo synthesis requirements for dihydrofolate and its precursors are small. Also, since exogeneous folates cannot be transported and used by *E. coli*, dihydrofolate synthesis is essential, but only at replenishment levels as a cell grows and divides. Because *pabA* and *pabB* are unlinked to one another and to other known *fol* genes (6, 8, 35), the question of how they are regulated arises. Before this question can be answered, however, more information on the organization of each particular gene and its expression signals is required.

The nucleotide sequences of *pabA* from *E. coli*, *Salmonella typhimurium*, and *Klebsiella aerogenes* were previously determined and were found to be preceded by a conserved open reading frame terminating 31 base pairs (bp) before the *pabA* initiation codon, suggesting that in these organisms *pabA* might be part of an operon (18). The studies presented here extended the structural analysis of the *pabA* region in *E. coli* and *S. typhimurium* and investigated the parameters affecting expression of *E. coli pabA*.

* Corresponding author.
† Present address: Betagen Corp., Waltham, MA 02154.

MATERIALS AND METHODS

Strains, media, and genetic techniques. All bacterial strains, plasmids, and phages are listed and described in Table 1 (also see Fig. 5).

E. coli cultures were grown at 37°C either in NZY broth (24) or in Vogel-Bonner minimal salt medium (41). Amino acids and thiamine, when needed, were supplemented at the concentrations recommended by Davis et al. (11); PABA was used at 10 ng/ml. Antibiotics were used at 50 μg/ml for ampicillin and kanamycin and at 5 μg/ml for chloramphenicol.

Transformation was carried out by the $CaCl_2$ method (24); P1 transduction was done by the method of Miller (29).

Chemicals and enzymes. PABA, amino acids, antibiotics, isopropyl-β-D-thiogalactoside, *o*-nitrophenyl-β-D-galactopyranoside (ONPG), and bovine serum albumin were purchased from Sigma Chemical Co. 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside was purchased from Research Organics. Synthetic 8-mer and 12-mer *Bgl*II linkers were purchased from New England BioLabs, Inc. Oligonucleotides BPN003 (5'-TACTGGTAGAGGTTCCAGGT) and BPN023 (5'-AGAACCGCGCTACTCAGATTCC) were synthesized by the Laboratory for Molecular Biology DNA Synthesis Facility at the University of Illinois at Chicago. Other oligomers used in DNA sequencing were synthesized or purchased from New England BioLabs. All enzymes used in standard DNA techniques and RNA analysis were purchased from Boehringer Mannheim Biochemicals, New England BioLabs, Promega Biotec, or United States Biochemical Corp. T4 DNA ligase was prepared in this laboratory by a previously published procedure. All radioactive nucleotides were purchased from Amersham Corp.

DNA manipulation. Standard recombinant DNA techniques, such as plasmid DNA isolation, restriction enzyme analysis, *Bal* 31 deletion, filling in of staggered DNA termini, linker ligation, agarose gel electrophoresis, and polyacryl-

TABLE 1. Bacterial strains, plasmids, and phages

E. coli strain, plasmid, or phage	Relevant properties[a]	Reference or source
Strains		
MC1000	Δ(*lacIPOZYA*)*X74*	37
MC4100	Δ(*argF-lac*)*U169 pro*$^+$	37
JC7623	*recB21 recC22 sbcB15 proA2*	44
JM103	Δ(*lac-pro*)(F′ *traD proAB lacI*^{q} *lacZM15*)	28
CJ236	*dut ung*(F′)	Amersham
W3110	*trpA33*	C. Yanofsky
N100	*galK*	27
BN138	JC7623 *pro*$^+$ (*argF-lac*)*U169*	This work
BN141	BN138 Φ(*pabA-lacZ*)*17*(Hyb)[b] *cam*	This work
BN158	MC1000 *PvuII*$_b$::*kan* Φ(*pabA-lacZ*)*34*(Hyb)[c] *cam*	This work
BN168	MC1000 *NruI*::*kan* Φ(*pabA-lacZ*)*34*(Hyb) *cam*	This work
BN184	MC1000 *BglII*::*kan* Δ*PvuII*$_a$-*PvuII*$_b$ Φ(*pabA-lacZ*)*34*(Hyb) *cam*	This work
BN185	MC1000 *BglII*:*kan* Δ*HpaI*-*PvuII*$_b$ Φ(*pabA-lacZ*)*34*(Hyb) *cam*	This work
BN186	MC1000 *BglII*::*kan* Φ(*pabA-lacZ*)*34*(Hyb) *cam*	This work
BN188	MC1000 *PstI*::*kan* Φ(*pabA-lacZ*)*34*(Hyb) *cam*	This work
BN193	MC1000 *PstI*::*kan* ΔP$_1$ Φ(*pabA-lacZ*)*34*(Hyb) *cam*	This work
BN198	MC1000 *BglII*::*kan* ΔP$_1$ Φ(*pabA-lacZ*)*34*(Hyb) *cam*	This work
BN501	MC1000 *BglII*::*kan*	This work
BN502	MC1000 *BglII*::*kan* ΔP$_1$	This work
BN503	MC1000 *PvuII*$_b$::*kan*	This work
Plasmids and phages		
pBN*pabA*	*E. coli pabA*	19
pJK8	*S. typhimurium pabA*	18
pKO-11	Promoterless *galK ap*	27
pMLB1034	Promoterless and truncated *lacZ ap*	42
pMC1871	Truncated *lacZ* cassette	Pharmacia
pMB2190	*kan* gene cassette	M. Casadaban
pMK2004	Cloning vector; *ap tet kan*	Laboratory collection
pBR322	Cloning vector; *ap tet*	6
pMH100	Contains *Bacillus subtilus cam* gene	M. Hulett
pT9	pBR322 containing *B. subtilus cam* cassette	This work
pBS	RNA probe-generating vector	Stratagene
M13mp8, -9, -18, -19	Templates for DNA sequencing and mutagenesis	28
P1*kc*	Generalized *E. coli* transducing phage	Laboratory collection

[a] Complete genotypes are not given.
[b] *lacZ* fused at *pabA* codon 17.
[c] *lacZ* fused at *pabA* codon 34.

amide gel electrophoresis were performed by the methods of Maniatis et al. (24).

Bacterial chromosomal DNA was prepared by the lysozyme-sodium dodecyl sulfate-phenol extraction method. Southern analysis was performed by the procedure supplied by Dupont, NEN Research Products, by using a GeneScreen Plus membrane. Nick-translated DNA with [α-^{32}P]dCTP was used for DNA hybridization.

DNA sequence analysis. The source of *E. coli* DNA fragments was pSZD3, which contains the *PvuII*$_a$-*PstI* fragment (Fig. 1A) cloned in pKO-11 (27; S. Z. Doktor and B. P. Nichols, unpublished data), and the source of *S. typhimurium* DNA fragments was pJK8 (18). DNA fragments were cloned into bacteriophage M13mp8, -9, -18, or -19 (28), and sequences were determined by the Sanger dideoxy-chain termination method (32) by using conditions described previously (18). The sequencing strategy is shown in Fig. 1. The nucleotide sequences of one fragment containing strong secondary structure-forming sequences could not be unambiguously resolved by substituting dITP for dGTP or by raising the DNA polymerase reaction temperature. The Maxam-Gilbert procedure (25) was used to confirm the nucleotide sequence of that fragment. The nucleotide sequence was determined on both strands of the DNA, and all restriction sites for cloning DNA fragments were confirmed by sequence determination of overlapping DNA fragments (Fig. 1 and 2).

RNA analysis. RNA was isolated from mid-log-phase cultures grown with aeration in NZY broth by the hot phenol method as modified by Arps et al. (3) and was followed by DNase treatment (RQ1 DNase; Promega Biotec) to remove residual contaminating DNA.

RNase protection mapping was performed by the method given in reference 5. The uniformly labeled sense or antisense RNA probes were transcribed from linearized pT48 template containing the *PstI-HincII* fragment of pBN*pabA* (Fig. 1A) cloned in pBS (Stratagene) in the presence of [α-^{32}P]CTP by using T3 or T7 RNA polymerase. Approximately 10^5 cpm of probe was hybridized to 25 μg of RNA or yeast tRNA (as control) in 30 μl of hybridization buffer (80% formamide, 40 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) [pH 6.4], 0.4 M NaCl, 1 mM EDTA [pH 8]) at 45°C for approximately 12 h. Subsequent RNase A and RNase T1 digestion was performed at 30°C for 90 min. Digestion products were electrophoresed on 5% polyacrylamide–7 M urea sequencing gels and analyzed by autoradiography. Sequencing reactions of M13mp18 were loaded on the same gel to approximate the sizes of protected fragments.

Nuclease S1 protection mapping was done as described by

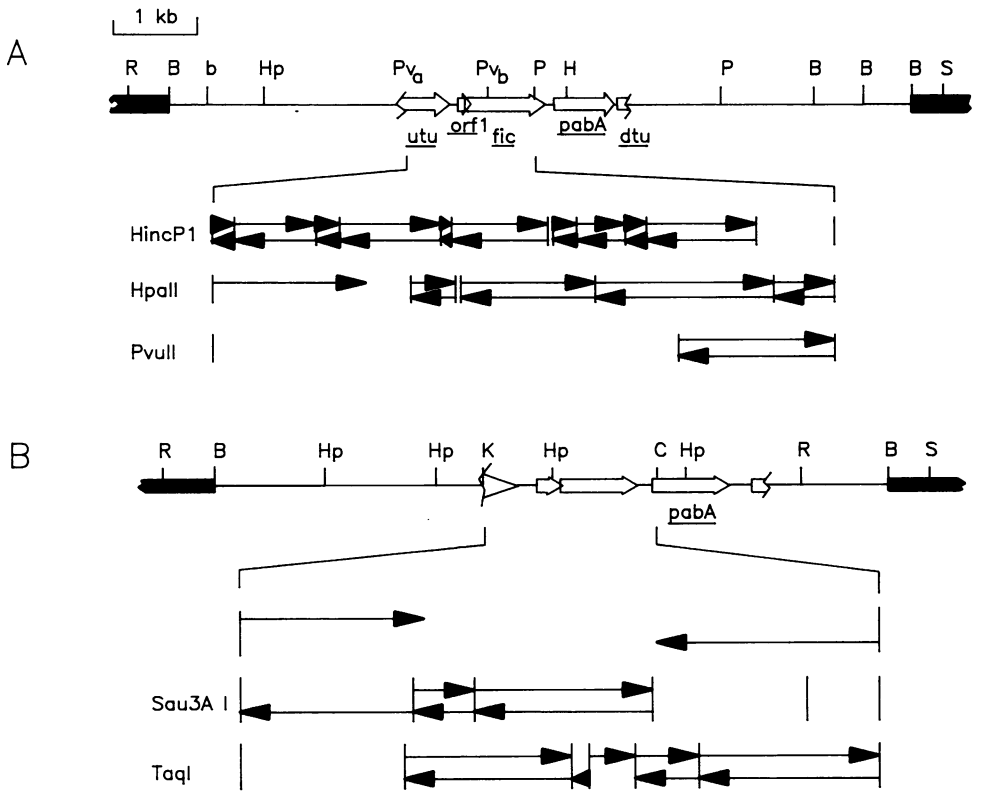

FIG. 1. Restriction maps of pBN*pabA* (A) and pJK8 (B) and sequencing strategies of the indicated regions. Thin and thick lines indicate chromosomal and plasmid DNA, respectively. Open arrows or boxes indicate coding regions deduced from DNA sequence analysis. *utu*, *orf1*, *fic*, and *dtu* are designated as discussed in Results. Vertical lines indicate positions of restriction sites, and thin arrows or arrowheads represent the direction and extent of sequence determined from the indicated regions. Restriction sites are indicated as follows: B, *Bam*HI; b, *Bgl*II; b/B, inactive *Bgl*II-*Bam*HI hybrid; C, *Cla*I; H, *Hin*cII; H/Sm, inactive *Hin*cII-*Sma*I hybrid; Hp, *Hpa*I; K, *Kpn*I; N, *Nru*I; P, *Pst*I; Pv, *Pvu*II; R, *Eco*RI; Sa, *Sac*I; S, *Sal*I; Sc, *Sca*I; Sm, *Sma*I; X, *Xho*I.

Arps et al. (3). Isolated DNA fragments were end labeled by using [γ-^{32}P]ATP and T4 polynucleotide kinase, and separated strands were prepared as described by Maxam and Gilbert (25). The labeled DNA was hybridized with 350 μg of RNA at 32°C for 15 h. Subsequent S1 digestion was performed at 15°C for 0 and 10 min, and the products were analyzed as described above.

RNA primer extension mapping of 5′ ends was performed essentially as described in reference 5. The oligonucleotide primer BPN003, which corresponds to nucleotides 1396 to 1415 within *pabA* (Fig. 2), was end labeled by using [γ-^{32}P]ATP and T4 polynucleotide kinase. The labeled primer and unincorporated nucleotides were fractionated on a NAP10 column (Sephadex G25; Pharmacia). The primer fraction was subsequently lyophilized and hybridized to 50 μg of RNA in 30 μl of hybridization buffer at 45°C for 12 h. Extension of the primer was carried out at 42°C for 90 min by using avian myeloblastosis virus reverse transcriptase (Promega Biotec), and the extension products were analyzed as described above. To locate the exact 5′ ends of the extended transcripts, plasmid pSZD13, which contains the *Pvu*II$_a$-*Hin*cII fragment (see below), was sequenced with the same primer and run along with the extension products.

Construction of *galK* transcriptional fusion plasmids. *galK* transcriptional fusions were constructed in either pKO-6 or pKO-11 vectors (27). pKO-6, pKO-11, and pKG1800 (containing the *gal* promoter) were obtained from M. Winkler. The source of insert DNA was pBN*pabA* (19), MPK1 (19), or M13 phages used for DNA sequence analysis. Figure 3 illustrates the extent of DNA present in each plasmid used in this study. Each plasmid has a number of intermediate constructions between its source DNA and the final product. Details of the constructions are available from the authors upon request.

pSZD2 contains DNA from the *Hpa*I site to the *Pst*I site in *fic* inserted into pKO-11. pSZD42 was derived from pSZD2 by deleting DNA between the *Pvu*II$_a$ site and the polylinker present between the insert DNA and *galK*. pSZD10 is a pKO-11 vector that carries the *Pst*I-*Hin*cII fragment containing the *fic-pabA* intercistronic region. pSZD12 was constructed from portions of pSZD2 and pSZD10 to reconstruct a pKO-11 vector containing DNA from the *Hpa*I site to the *Hin*cII site in *pabA*. pSZD6 contains the *pabA-dtu* intercistronic region and was constructed by ligating a 177-bp *Sau*3AI fragment derived from MPK1 with *Bam*HI-digested pKO-11. pSZD15 was constructed from an M13mp9 phage used for DNA sequence analysis and carries a 596-bp *Taq*I DNA fragment containing the *utu-orf1* intercistronic region inserted into pKO-6.

Construction of *pabA-lacZ* translational fusions in the *E. coli* chromosome. The *pabA-lacZ* translational fusion was constructed by first subcloning the *Bam*HI-*Pst*I fragment from pBN*pabA* into pMK2004 to yield pT2 (Fig. 4A). Second, the unique *Sca*I site of pT2 (in codon 18 of *pabA*) was converted to a *Bgl*II site by the addition of a synthetic 8-mer *Bgl*II linker to generate plasmid pT3. Plasmid pT3 was then

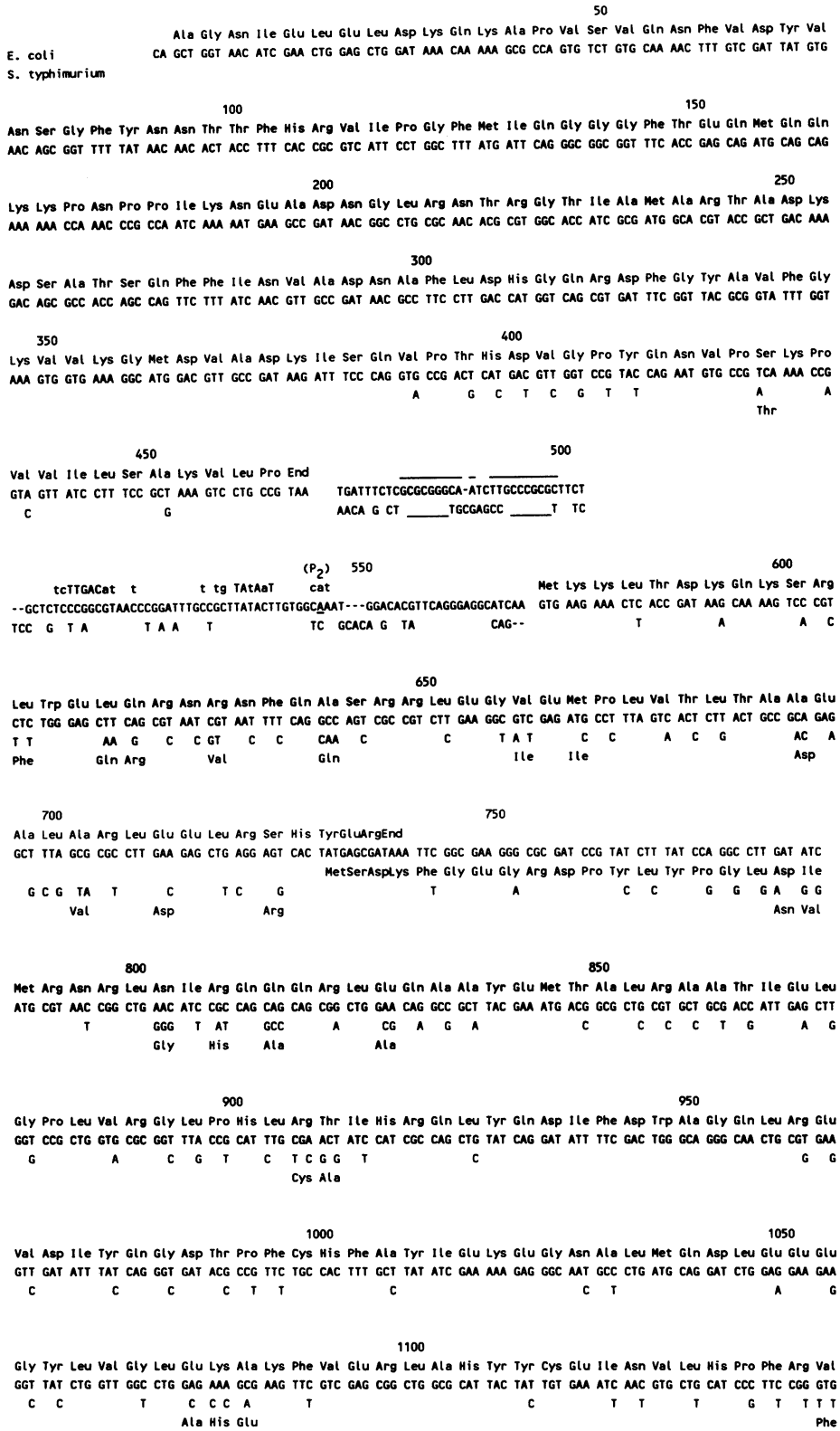

FIG. 2. DNA sequence comparison of the *utu-fic-orf1-pabA-dtu* regions from *E. coli* and *S. typhimurium*. The complete nucleotide and amino acid sequences of *E. coli* are presented, and only differences found in *S. typhimurium* are shown. Numbering is from the first residue of the *E. coli* *Pvu*II$_a$ site and applies only to the *E. coli* sequence. Hyphens represent gaps inserted to increase identities in the alignment. Upper- and lowercase nucleotides above the *E. coli* sequence show the consensus promoter sequence of Hawley and McClure (13). The lines on top of the sequence indicate elements of dyad symmetry, and the underlined nucleotides indicate the 5′ terminus of the transcript initiated by the indicated P_1 and P_2 promoters. The *pabA* and *dtu* sequences were previously published (18).

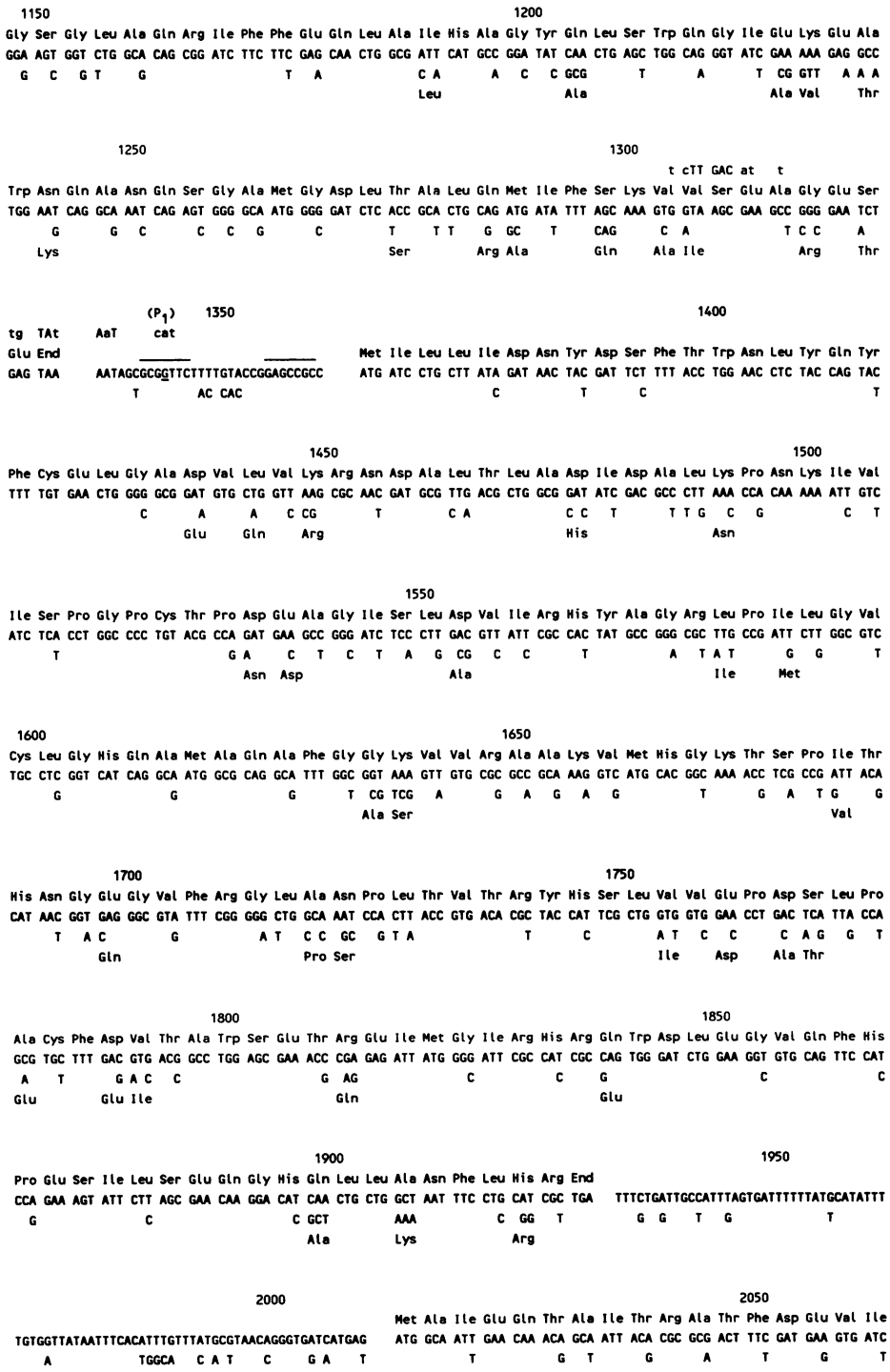

FIG. 2—*Continued*

digested with *Bgl*II and ligated to the *Bam*HI *lacZ* fragment excised from pMC1871 (Pharmacia) and to the *Bam*HI *cam* fragment from pT9. DNA from Kan^r Cam^r blue (on 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside indicator plate) transformants was isolated for restriction analysis. The correct construct was designated pT12.

Plasmid pT12 was linearized at the *Xho*I site (in the *kan* gene) and transformed into strain BN138 (*recBC sbcB* Δ*lac*). Cam^r blue transformants were selected and screened for the Kan^s and PABA$^-$ phenotype, suggesting the absence of plasmid and homologous exchange at the *pabA* locus (44). The chromosomal organization of the transformant designated BN141 was verified by Southern analysis (data not shown).

Construction of deletions and/or *kan* insertions in the upstream region of the *pabA-lacZ* fusion. A second *pabA-lacZ* translational fusion was derived from pSZD11. Plasmid pSZD11 is a derivative of pKO-11 containing the 3.6-kilo-

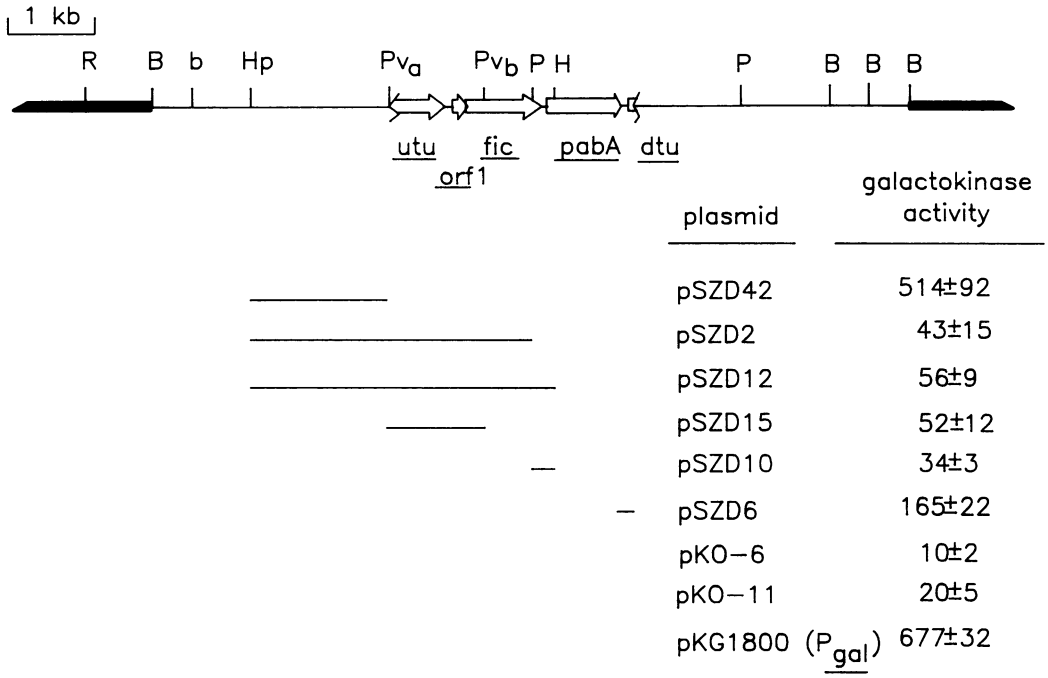

FIG. 3. Analysis of promoter activity in *galK* transcriptional fusions. The physical map of the region near *pabA* cloned in pBN*pabA* is shown at the top, and segments cloned into promoterless *galK* vectors are illustrated below. Thin and thick lines indicate chromosomal and plasmid DNA, respectively. pSZD15 is a pKO-6 derivative, and the remainder are pKO-11 derivatives. Restriction sites are abbreviated as described in the legend to Fig. 1. Units of galactokinase are expressed as micromoles of galactose-1-phosphate per nanomole of plasmid. Values are averages of six determinations.

base *Bam*HI-*Hin*cII fragment flanked by two *Eco*RI sites, one in the vector upstream of the insert and the other within the polylinker where *pabA* is transcriptionally fused to *galK* (Fig. 4B). The *Eco*RI insert was excised from plasmid pSZD11 and cloned into plasmid pMLB1034 at the *Eco*RI site, resulting in an in-phase fusion of *pabA* at codon 34 to *lacZ* at codon 9. This construct was designated pSZD21 and was used to generate the *kan* insertion derivatives.

The *kan* marker was inserted upstream of *pabA* at the *Bgl*II, *Nru*I, and *Pst*I sites (Fig. 5) as follows. The *Bgl*II::*kan* construct was made by digesting pSZD21 with *Bgl*II and ligating it to the *kan Bam*HI fragment excised from pMB2190 (pT35, Fig. 5). The *Nru*I::*kan* construct was made in several steps. First, the *Bgl*II site of pSZD21 was abolished by *Bgl*II digestion followed by filling in of the ends with DNA polymerase I (Klenow fragment) and dNTPs and religating. Second, the resulting construct was digested with *Nru*I and ligated to the 8-mer *Bgl*II linker to place a convenient cloning site at that position. Finally, the *kan Bam*HI fragment from pMB2190 was inserted at the new *Bgl*II site to generate pT18 (Fig. 5). The *Pst*I::*kan* construct was made by partially digesting pSZD21 with *Pst*I and ligating it to the *kan Pst*I fragment excised from pMB2190 (pT40, Fig. 5).

The *Pvu*II$_b$::*kan* construct was made as follows. (i) Plasmid pSZD13 (which is the same as pSZD11 except that the *Eco*RI insert contains instead the *Pvu*II$_a$-*Hin*cII region and the *Pvu*II$_a$ site has been destroyed) was linearized at the *Pvu*II$_b$ site and ligated to the 12-mer *Bgl*II linker. (ii) The *Eco*RI insert of this construct was excised and cloned into pMLB1034 at the *Eco*RI site to generate the same *pabA-lacZ* fusion as in pSZD21 and derivatives (see above). (iii) The resulting construct was digested with *Bgl*II and ligated to the *kan Bam*HI fragment to generate pT22 (i.e., *Pvu*II$_b$::*kan*, Fig. 5).

Deletion of the *Hpa*I-*Pvu*II$_b$ region in the presence of the *Bgl*II-*kan* marker was constructed as follows. (i) Plasmid pSZD11 was digested with *Hpa*I and *Pvu*II and religated. (ii) The *Hpa*I-*Pvu*II$_b$ deletion construct was digested with *Bgl*II and ligated to the *kan Bam*HI fragment. (iii) The *Eco*RI insert of the new construct now containing the *Bgl*II::*kan* and *Hpa*I-*Pvu*II$_b$ deletion was excised and cloned into plasmid pMLB1034 at the *Eco*RI site to generate the same *pabA-lacZ* fusion as in the constructs described above. This construct was designated pT38 (Fig. 5).

The *Pvu*II$_a$-*Pvu*II$_b$ deletion in the presence of the *Bgl*II::*kan* marker was constructed by first digesting pSZD11 with *Pvu*II and religating; the construct was then digested with *Bgl*II and ligated to the *kan Bam*HI fragment. The out-of-phase fusion between *utu* and *fic* resulting from the deletion between the two *Pvu*II sites was made in phase by introducing an 8-mer *Bgl*II linker at the *Pvu*II$_a$-*Pvu*II$_b$ junction. The *Eco*RI insert containing the newly constructed region was excised and cloned into pMLB1034 to generate pT37 (Fig. 5).

The structures of these constructs and their intermediates were verified by restriction and/or DNA sequence analyses, and the orientation of the inserted *kan* marker was determined by restriction analysis.

All of the above constructs were crossed into the *E. coli* chromosome by linearizing the plasmids at the *Sac*I site and transforming them into strain BN141. Crossing of these constructs into the chromosome was facilitated by homologous recombination between the regions upstream of the *kan* insertion and downstream of the *lacZ* portion of the fusions present in the chromosome and plasmid (Fig. 4B). Transformants were selected for Kan^r and screened for Ap^s, which suggested the absence of plasmids. Since the *pabA-lacZ* fusions in strain BN141 and in plasmid constructs were not the same (i.e., *lacZ* was fused at codons 17 and 34 of *pabA*, respectively [Fig. 4]), Southern analysis was carried out to

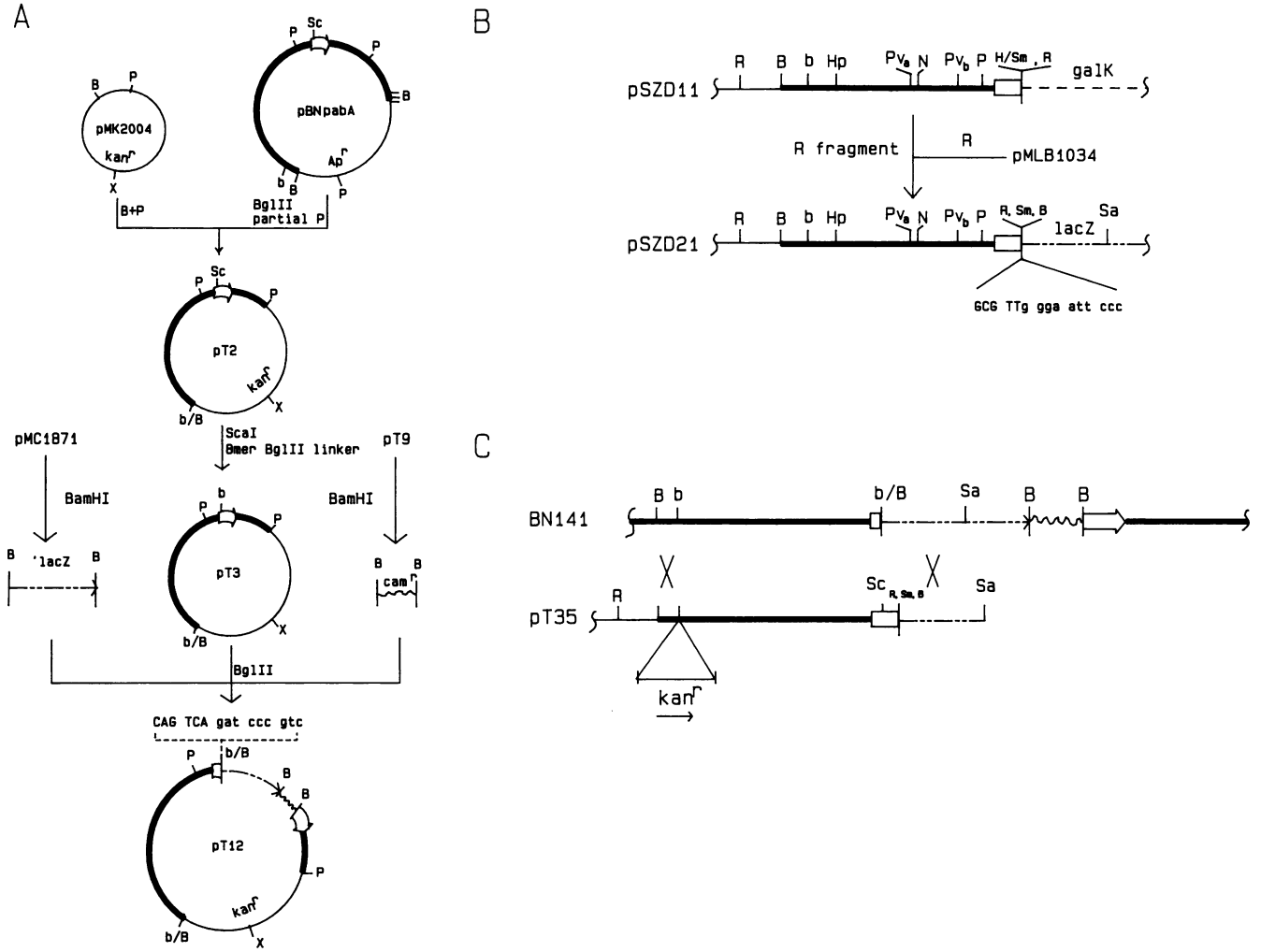

FIG. 4. Construction of *lacZ* protein fusions at codons 17 and 34 of *pabA* and strategy for crossing plasmid constructs into the *E. coli* chromosome by homologous recombination. (A) Construction and structure of the codon 17 *pabA-lacZ* protein fusion. (B) Construction and structure of the codon 34 *pabA-lacZ* protein fusion. (C) Crossing of the codon 34 *pabA-lacZ* fusion and the *Bgl*II::*kan* marker (pT35) into strain BN141. Cross marks indicate the region of the double crossover event between the chromosome and plasmid that generates a chromosomal *Bgl*II::*kan* Φ(*pabA-lacZ*)*34*(Hyb) strain. kan^r, ap^r, and cam^r indicate markers for resistance to kanamycin, ampicillin, and chloramphenicol, respectively. Symbols: □, *pabA*-coding region; →, *kan* marker; ▬, chromosomal sequence flanking *pabA*; ——, vector. Nucleotides written in capital letters represent part of the *pabA* and linker sequence, and those in lowercase letters represent part of the *lacZ* sequence. Restriction sites are abbreviated as described in the legend to Fig. 1.

screen for transformants containing the codon 34 fusion (data not shown). This was facilitated by the presence of a *Bam*HI site at the junction of the codon 34 fusion. Only transformants carrying the codon 34 fusion were subsequently transduced into strain MC1000 for expression analysis. These strains are illustrated in Fig. 5.

Oligonucleotide-directed mutagenesis. Uracil-containing DNA template of phage mpt4 (see Fig. 8) was prepared from *E. coli* CJ236 (*dut ung*) and mutagenized (5) by using the primer BPN023. The mutation was verified by DNA sequence analysis. The insert from the mutant phage designated mpt11 was excised with *Sma*I and cloned into *Sma*I-digested pT40*Bal*213 (see Fig. 8).

Enzyme assays. β-Galactosidase activity was measured as described by Miller (29). Twenty-five-milliliter mid-log-phase cultures were harvested, washed with 0.85% NaCl, and suspended in 1 ml of Z buffer. The cell suspension was sonicated for 30 s at 50% pulse, cooled in ice for 30 s, and sonicated for an additional 30 s. Centrifugation was carried out to remove the cell debris from the lysate. The lysate was immediately assayed for protein concentration by using the Bio-Rad method with bovine serum albumin as a standard and for β-galactosidase activity by using ONPG as the substrate. One unit of β-galactosidase is defined as 1 nmol of ONPG hydrolyzed per min per mg of protein. An extinction coefficient for ONPG of 4,500 was used to calculate moles of ONPG (37).

Galactokinase assays were carried out as described by McKenney et al. (27), with the modifications suggested by Jones et al. (17) incorporated. Plasmid copy number was determined by the hybridization method described by Adams and Hatfield (1), and 1 U of galactokinase activity is defined as 1 μmol of galactose-1-phosphate per nmol of plasmid.

Growth rate assay. Cells were grown overnight in Vogel-Bonner minimal medium supplemented with leucine, thiamine, 0.2% glucose, and PABA (10 ng/ml) and diluted 1:100 in 10 ml of the same medium with or without PABA

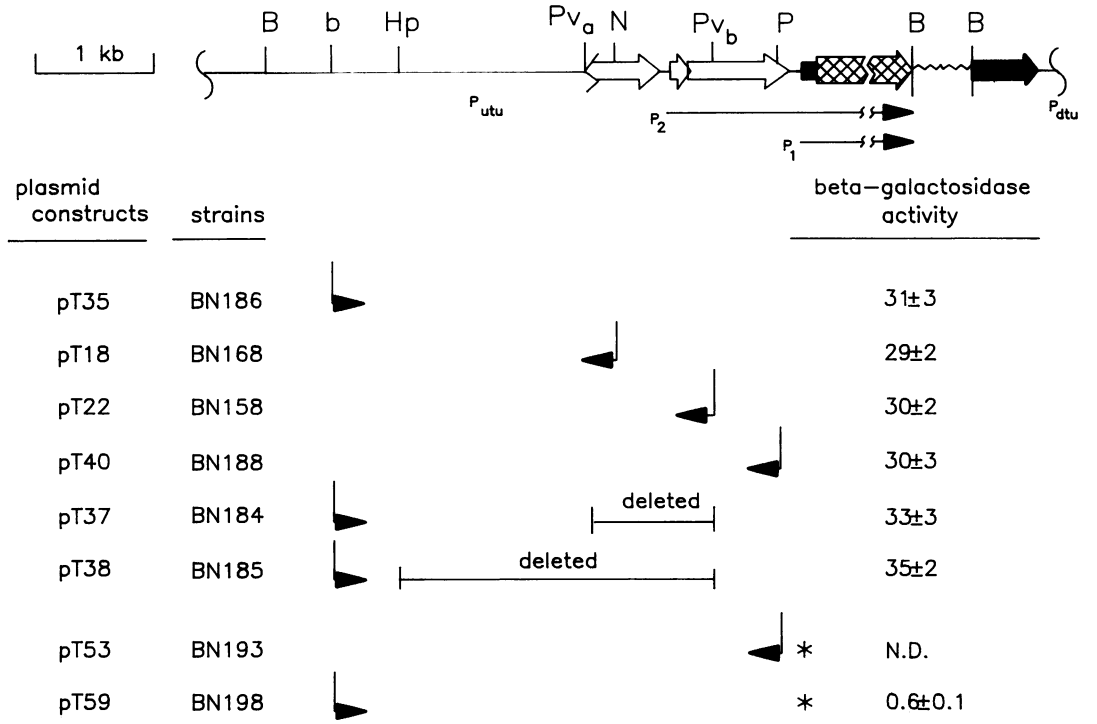

FIG. 5. Chromosomal organization and expression of *pabA-lacZ* protein fusion and deletion and/or *kan* insertion derivatives. The physical and genetic map is shown at the top. Restriction sites are abbreviated as described in the legend to Fig. 1. Portions of *pabA* (■) and *lacZ* ( ) in the fusion, the *cam* marker (~~~), and the 3′ region of *pabA* ( ) are shown. Transcriptional units including *pabA* are indicated below the chromosomal map; also shown are the regions containing promoters for *utu* (P_u) and for *dtu* (P_d). Thin arrows represent the position and orientation of the *kan* marker in corresponding plasmid constructs or chromosomal strains designated at the left, and asterisks represent the P_1 deletion mutation (see Fig. 8) in the indicated strains. Units of β-galactosidase are expressed as nanomoles of ONPG per minute per milligram of protein. N.D., Not detected. Values are averages of four determinations.

supplementation. The cultures were shaken at 37°C, and turbidity was determined with a Klett-Summerson meter at 30-min intervals after the first 3 h.

RESULTS

Nucleotide sequence upstream of *pabA*. We have previously noted the presence of an open reading frame which ends 31 bp upstream of the *pabA* initiation codon (18). This suggested that *pabA* might be organized into an operon including the upstream gene(s). Nucleotide sequence analysis was therefore undertaken to investigate the structure of the putative transcriptional unit encoding *pabA*. Mini-Mu dl (*lacZ*; D. Dix and B. P. Nichols, unpublished data) and *galK* transcriptional fusions (see below) indicated that the 5′ end of a transcriptional unit lay 1.2 to 2 kilobases upstream of *pabA*. The target chosen for sequence analysis was therefore the region between the *Pvu*II$_a$ site and the coding region of *pabA* (Fig. 1A). We also determined the nucleotide sequence of an equivalent region from *S. typhimurium*. The strategies for determining the nucleotide sequences are shown in Fig. 1.

The nucleotide sequence comparison of the 2,062-bp region including *pabA* and downstream sequences (taken from reference 18) is shown in Fig. 2. Nucleotide difference patterns between *E. coli* and *S. typhimurium* suggested the presence of four protein-coding regions in addition to *pabA*; only one of these genes, designated *fic*, has been identified, and it is thought to encode a product involved in cell division (20, 21).

The *Pvu*II$_a$ site appears to lie within the unidentified upstream transcriptional unit (*utu*) which is separated by 104 bp from the beginning of another unidentified open reading frame, designated *orf1*. Within this intergenic region is a putative ρ-independent transcription terminator which may serve to prevent transcription readthrough beyond *utu*. *orf1* encodes a polypeptide of 55 amino acids in length with a mass of 6,550 daltons. *orf1* overlaps a Shine-Dalgarno sequence (36), an ATG triplet, and 8 bp (not including the termination codon) of another open reading frame, identified as *fic*. The *fic*-coding region extends for 200 codons (600 bp) and encodes a 22,950-dalton protein. This gene lies 31 bp upstream of *pabA* and was identified by comparison of restriction site maps reported for *fic* (20). An open reading frame 85 bp downstream from *pabA* that is conserved in *S. typhimurium* and *K. aerogenes* signifies the beginning of another gene, which belongs to the designated downstream transcriptional unit (*dtu* [18]) and likely encodes *argD* (20, 21).

The specific translational phases predicted from DNA sequence analysis through the *orf1*, *fic*, and *pabA* region were confirmed by constructing *lacZ* translational fusions by using the plasmid vector pMLB1034 (42). DNA fragments derived from phage M13mp8 subclones were used to make fusions with the reading frames predicted in Fig. 2 at three points within *orf1* (at nucleotides 613, 660, and 691), three points within *fic* (at nucleotides 756, 782, and 923), and one point in *pabA* at nucleotide 1463 (pSZD21; see below). The plasmids conferred a Lac$^+$ phenotype in *E. coli* MC1000 as detected by a blue color on 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside indicator plates. The production of β-galactosidase indicated that translation through *orf1*, *fic*,

and *pabA* occurred in the phases predicted by nucleotide sequence analysis.

Recently, *fic* and the upstream region were completely sequenced by Kawamukai and co-workers (21). The sequence of *fic* reported is identical to that presented here. Moreover, *utu* was found to encode a single open reading frame of 190 amino acids. *orf1*, however, was not identified by these workers, although the nucleotide sequence reported is identical to that presented here.

Location of promoters in the vicinity of *E. coli pabA*. Promoters in the vicinity of *pabA* were localized by constructing a series of *galK* transcriptional fusions, using the promoterless *galK* vectors pKO-6 and pKO-11 (27). The levels of galactokinase expression were corrected for gene dosage effects (1) and compared with those directed by pKO-6 and pKO-11 (containing no promoter) and pKG1800 (containing the *galK* promoter). The data are summarized in Fig. 3.

A strong promoter that lay between the *Hpa*I and *Pvu*II$_a$ site (P_{utu}) was detected that directed transcription through the *Pvu*II$_a$ site in *utu* (pSZD42, Fig. 3). P_{utu} directed galactokinase expression at levels slightly less than pKG1800, which contained the *galK* promoter. Plasmids that contained P_{utu} and that had 3′ endpoints at the *Pst*I (within *fic*) or *Hin*cII (within *pabA*) sites expressed about 10% as much galactokinase as did pSZD42, suggesting that P_{utu} transcription was terminated before the *pabA*-coding region. Insertion of DNA fragments that spanned the *utu-orf1* and *fic-pabA* intercistronic regions (pSZD15 and pSZD10, respectively) resulted in roughly similar levels of galactokinase expression, although the levels were about 10% of that expressed from P_{utu}. A promoter downstream of *pabA* (P_{dtu}) was localized to a 177-bp *Sau*3AI DNA fragment containing the last 81 nucleotides of *pabA* and ending 5 bp before the *dtu* initiation codon. Plasmids containing P_{dtu} (e.g., pSZD6) expressed galactokinase at approximately 33% of the level of P_{utu}. The transcriptional analysis indicated that the *orf1-fic-pabA* region was transcribed at low levels and lay between two other transcriptional units expressed at 10- and 3-fold-higher levels.

Expression analysis of *E. coli pabA*. To analyze regions involved in the expression of *pabA*, a chromosomal *pabA-lacZ* translational fusion and derivatives containing deletions and/or *kan* insertions in the upstream region of *pabA* were constructed in *E. coli* at the *pabA* locus by homologous recombination (44). The *kan* gene cassette from plasmid pMB2190 was used as a transcriptional block, since it exerts strong polarity on the expression of downstream genes if these genes are cotranscribed (4). Furthermore, it provides a convenient selection marker for crossing flanking homologous regions from plasmids into the *E. coli* chromosome. The *kan* cassette was inserted in the upstream region farthest from *pabA* (*Bgl*II), within *utu* (*Nru*I) and within *fic* (*Pvu*II$_b$ and *Pst*I). In addition, deletions removing *utu* and the 5′ region of *fic* (*Hpa*I-*Pvu*II$_b$) and between *utu* and *fic* (*Pvu*II$_a$-*Pvu*II$_b$), each containing the *kan* cassette at the *Bgl*II site, were constructed (see Materials and Methods).

β-Galactosidase assays of the *lac* fusion strains indicated that PabA expression was not affected by the perturbation of the upstream region (Fig. 5). That is, PabA synthesis, monitored by the activity of the PabA-LacZ fusion, remained virtually the same regardless of the insertions in or deletions of the region upstream of the *Pst*I site. These data indicated that *pabA* was expressed independently of any upstream genes and that the putative *pabA* promoter lay in the 82-bp region between the *Pst*I site and *pabA*. The initiation site of the promoter near *pabA* was determined by an RNase protection assay. The *Pst*I-*Hin*cII fragment (nucleotides 1282 to 1463), which spans the *fic-pabA* intergenic region, was cloned into the riboprobe plasmid pBS (Stratagene) and used to generate a ^{32}P-labeled antisense RNA probe. Figure 6A shows that all strains tested showed a protected RNA of about 127 nucleotides in length. The 5′ terminus of the transcript was precisely mapped by RNA primer extension by using a primer (BPN003) corresponding to codons 12 to 18 of *pabA*. The results, shown in Fig. 6B, indicated that the primer was extended to the guanine nucleotide at position 1341, which is 22 bases upstream of the *pabA* initiation codon (Fig. 2). This signal was generated from all strains and thus is in agreement with the results of the RNase protection assay.

Since the initiation site is only 10 bases from the *fic* stop codon (TAA), the deduced −10 and −35 regions of the *pabA* promoter (designated P_1) lie between positions 1304 and 1340 and thus within the *fic*-coding region and include the *fic* stop codon (Fig. 2 and 6B).

In addition to the transcript initiated by P_1, both RNase protection and RNA primer extension mapping indicated the presence of a transcript initiated upstream of the *Pst*I site (Fig. 6). The RNase protection assay showed full-length protection of the *fic-pabA* intergenic probe by RNA prepared from the *Bgl*II::*kan*, *Nru*I::*kan*, and Δ*Pvu*II$_a$-*Pvu*II$_b$ strains. Full-length protection was not observed with the *Pst*I::*kan* and Δ*Hpa*I-*Pvu*II$_b$ strains, suggesting that another message containing the intercistronic region was likely initiated from the region between the *Nru*I and *Pvu*II$_b$ sites, or more specifically, from the *utu-orf1* intergenic region. This region showed transcription activity in the *galK* fusion assay (pSZD15 in Fig. 3) and contains a structure resembling a promoter sequence (Fig. 2). The putative promoter lies between positions 507 and 545 and contains five identities to the −10 region of the *E. coli* promoter consensus sequence (12). S1 nuclease analysis was carried out to map the 5′ terminus of the transcript. The 5′-end-labeled 275-nucleotide *Hpa*II fragment (positions 520 to 795) which spans the *utu-orf1* intergenic region was used as a probe. The results, shown in Fig. 7, indicated the 5′ end of a transcript near position 545 in the *utu-orf1* intergenic region. The promoter initiating this transcript was designated P_2. Since little or no signal of full-length protection was observed, transcription from the *utu* region was probably terminated within the intercistronic region, possibly at the structure resembling a ρ-independent transcription terminator (nucleotides 477 to 502). Additional protection experiments using probes spanning the *orf1-fic* junction (positions 724 to 934) and within *fic* (*Pvu*II$_b$-*Pst*I) showed full-length protection (data not shown), indicating that these regions contain no promoter activity.

The relative steady-state levels of *pabA* transcripts initiated by P_1 and P_2 were determined by quantitating the signals that corresponded to full-length (P_2) and to partial (P_1) protection of the probe from the RNase mapping shown in Fig. 6A, lane 5. After the counts in each fragment were normalized to account for the fact that the full-length signal contains more labeled nucleotides, the data indicated that the levels of the two transcripts were approximately equal.

Transcription analysis therefore indicated that there are two transcriptional units that transcribe *pabA*: (i) the *pabA* monocistronic transcript initiated by P_1 and (ii) the *orf1-fic-pabA* polycistronic transcript initiated by P_2. Although the relative steady-state levels of *pabA* in the two units appeared the same, the translational study using *pabA-lacZ*

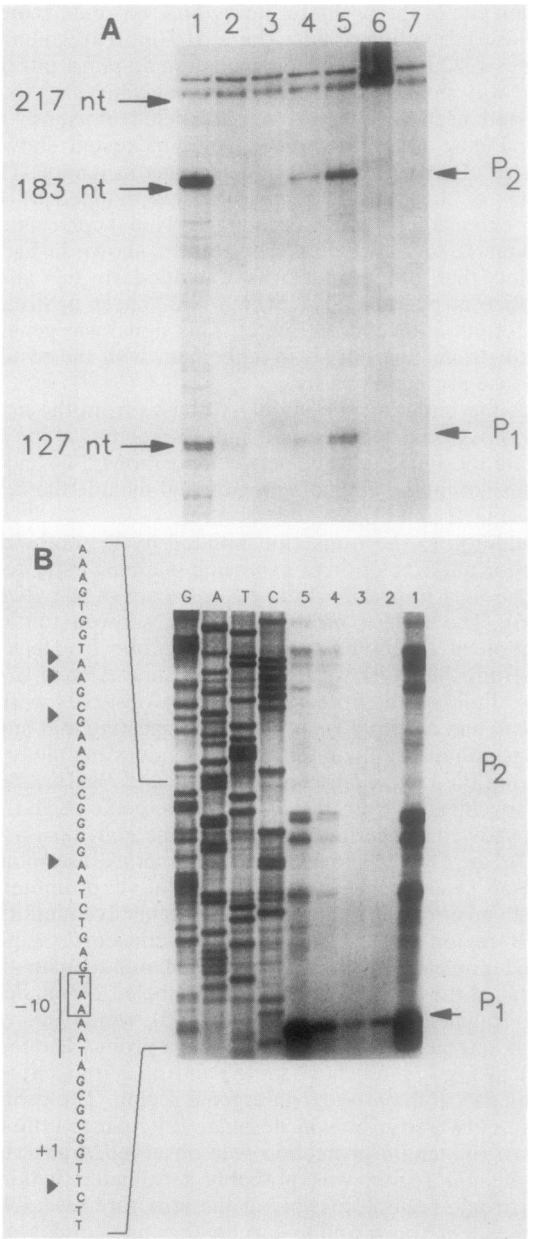

FIG. 6. Analysis of *pabA* transcripts by RNase protection and RNA primer extension mappings. (A) RNase protection. The antisense uniformly ^{32}P-labeled 217-nucleotide (nt) probe containing 183 nucleotides of the *Pst*I-*Hin*cII region (Fig. 1A), which spans the *fic* and *pabA* coding and intergenic region, and 34 nucleotides of the polylinker in plasmid pBS was transcribed by T7 RNA polymerase, hybridized to RNA prepared from the indicated strains, and digested with RNase A and T1. Lanes: 1, BN184 (*Bgl*II::*kan* Δ*Pvu*II$_a$-*Pvu*II$_b$); 2, BN185 (*Bgl*II::*kan*, Δ*Hpa*I-*Pvu*II$_b$); 3, BN188 (*Pst*I::*kan*); 4, BN168 (*Nru*I::*kan*); 5, BN186 (*Bgl*II::*kan*); 6, probe; 7, *S. cerevisiae* tRNA. The approximate sizes of the probe and of the protected fragments were determined from the M13mp18 sequencing ladder (not shown). P_1 and P_2 indicate the promoters initiating transcripts that generated the corresponding protected fragments. (B) RNA primer extension. The *pabA* primer which is complementary to nucleotides of codons 12 to 18 (positions 1396 to 1415) was hybridized to RNA prepared from the indicated strains and extended with avian myeloblastosis virus reverse transcriptase. Lanes 1 to 5, RNA prepared from strains as indicated for panel A; lanes G, A, T, and C, dideoxy sequencing ladder of plasmid pSZD13 (see Materials and Methods) with the same primer. Indicated on the left is the nucleotide sequence read from the sequencing ladder. The TAA stop codon of *fic* is boxed. The nucleotide of the 5′ end of the P_1-initiated transcript is indicated (+1). The deduced −10 region of P_1 is underlined. Arrowheads indicate the 5′ ends of the degraded products of the transcripts initiated by P_2 (see Discussion).

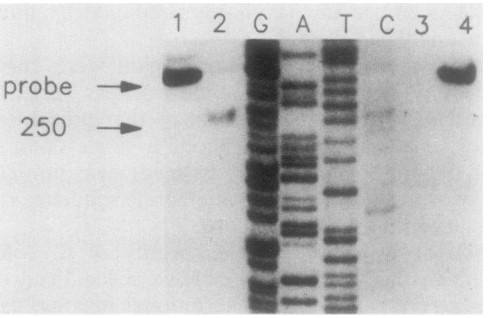

FIG. 7. Mapping of the initiation site of P_2 by S1 nuclease protection. The ^{32}P-labeled 275-nucleotide *Hpa*II fragment spanning the *utu-orf1* intergenic region was used as the probe. Lanes 1 and 2, Sense-strand probe hybridized to RNA of strain W3110 (0- and 10-min S1 nuclease digestion, respectively); lanes G, A, T, C, dideoxy sequencing ladder of mTAB8 used as size standards; lanes 3 and 4, antisense strand probe hybridized to RNA (10- and 0-min S1 nuclease digestion, respectively). The undigested probe and the size of the protected fragment (P_2) are indicated by arrows.

fusions indicated that PabA was expressed almost exclusively from the P_1-initiated transcript (Fig. 5).

Expression of PabA from the P_2-initiated polycistronic transcript. To test the role of P_2 in PabA expression in the absence of P_1, P_1 promoter activity was abolished by deleting the five bases of its −10 region (TAAAA, positions 1329 to 1333) by using oligonucleotide-directed mutagenesis. The deletion also removed the *fic* stop codon (TAA) since it is part of the −10 region. Translation termination of Fic is not affected since there is an in-phase nonsense triplet (TAG) also generated by the deletion (Fig. 8). The strategies of the mutagenesis using a single-stranded phage template and crossing the mutation into a plasmid containing the *pabA-lacZ* fusion are shown in Fig. 8. Plasmid pT40*Bal*213 was designed to facilitate the cloning of the mutated insert back into the *pabA-lacZ* fusion background by utilizing the same *Sma*I sites present on both the recombinant phage and the plasmid (Fig. 8). Since the *Sma*I insert contained the 5′-coding region of the *kan* gene and the *pabA* portion of the *pabA-lacZ* fusion and the *Sma*I-digested pT40*Bal*213 contained the 3′-coding region of the former and the *lacZ* portion of the latter, only recombinant plasmids resulting in reconstitution of the active *kan* gene (and thus the *lac* fusion) would be selected. The P_1 deletion and *kan* gene inserted at the *Bgl*II (pT59) or at the *Pst*I (pT53) site were then crossed into the *E. coli* chromosome as described above.

As expected, when P_1- and P_2-initiated transcription into *pabA* was abolished (by P_1 deletion and *kan* polar insertion, respectively), no β-galactosidase activity was detected (Fig. 5, strain BN193). Removal of the *kan* block, thus allowing P_2 transcription into *pabA*, did not result in an appreciable increase of β-galactosidase activity (Fig. 5, strain BN198). This indicated that P_2 expresses PabA at less than 2% of the level of P_1. RNA primer extension analysis detected only the P_2-initiated transcript in strain BN198 and no transcript in strain BN193 (data not shown). To confirm that the extension product generated by strain BN198 was not due to

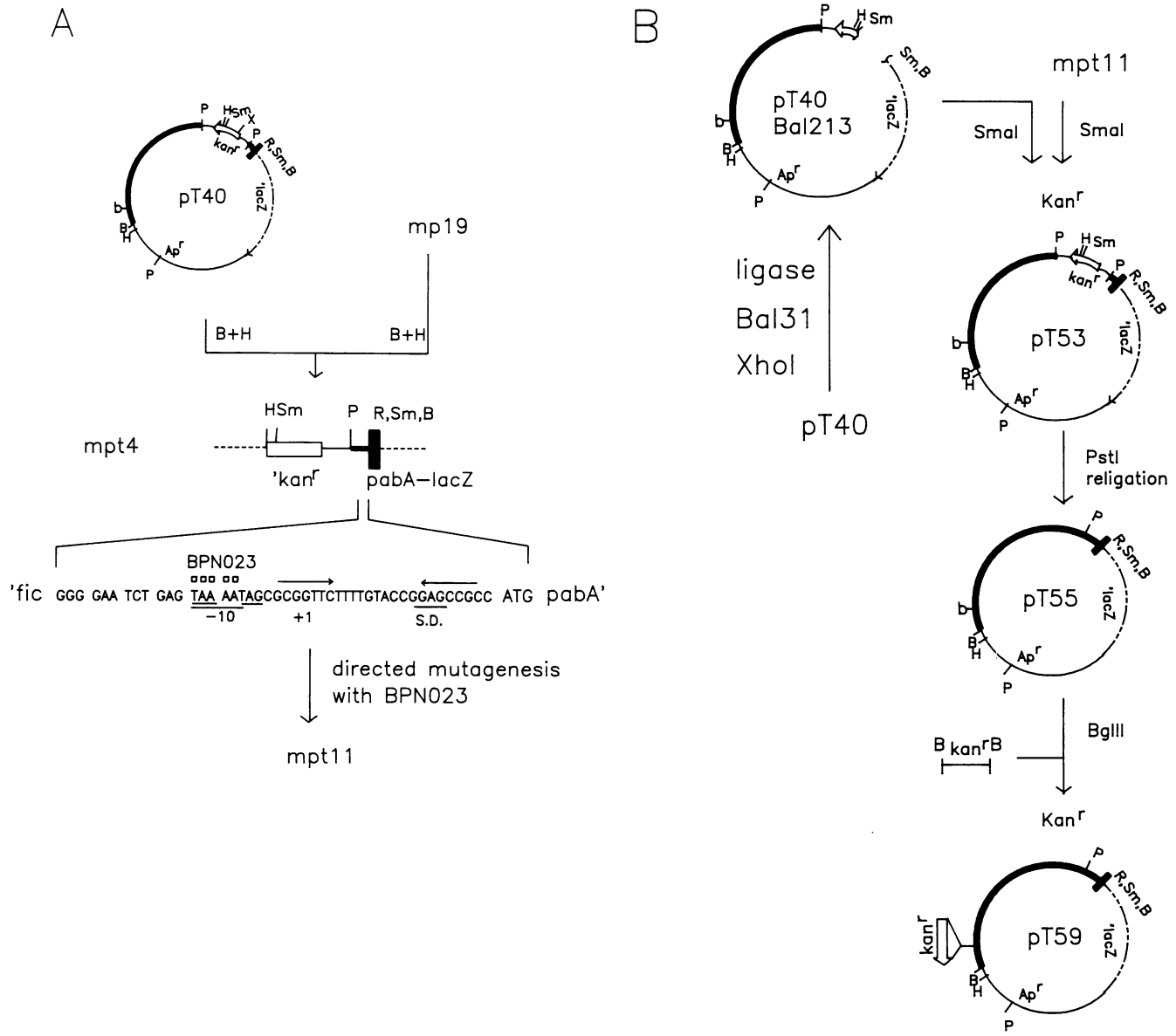

FIG. 8. Strategy of P_1 deletion mutation constructed by oligonucleotide-directed mutagenesis (A) and reconstitution of the *pabA-lacZ* fusion along with the mutation on plasmids (B). (A) mpt4 contains the 5′ region of the *kan* marker and the *pabA* portion of the *pabA-lacZ* fusion. The nucleotide sequence of the *fic-pabA* intergenic region is shown. Indicated on top of the sequence are the five bases in the −10 region of P_1 to be deleted (□) with the primer BPN023. Arrows on top of the sequence indicate elements of dyad symmetry. The natural (TAA) and artificial (TAG) stop codons of *fic* are underlined. +1, Transcription initiation site of P_1; S.D., Shine-Dalgarno sequence of *pabA*. (B) Strategy of reconstitution of the region containing the *kan* marker, the P_1 deletion, and the *pabA-lacZ* fusion onto plasmids. Abbreviations for restriction sites are given in the legend to Fig. 1.

residual DNA contamination in the RNA preparation, an RNase protection assay using the labeled sense RNA probe spanning the *fic-pabA* intergenic region was carried out. No protection of the probe was detected (data not shown), indicating that the primer extension product of strain BN198 was entirely due to the P_2-initiated transcript.

The results above indicated that P_2 expresses PabA at a very low level, which may or may not support *pabA*-dependent growth. In order to test this, the P_1 deletion was crossed from strain BN198 into a strain containing the wild-type *pabA*-coding region by phage P1 transduction. Transductants were first selected for Kan^r and screened for Cam^s, indicating that the *pabA-lacZ* region in the transducing DNA fragment was not crossed into the recipient chromosome. Kan^r Cam^s transductants were then screened for growth on minimal media lacking PABA. Several transductants grew as well as the wild-type strain, suggesting that the second crossover generating these transductants occurred upstream of the P_1 promoter region. One of these was designated strain BN501. Rare transductants, occurring at a frequency of about 2.5%, grew much more slowly than the isogenic wild type or strain BN501. This transductant, designated BN502, should contain the P_1 deletion mutation which resulted from the second crossover occurring in the region between the P_1 locus and codon 34 of *pabA*. The chromosomal organization at the *pabA* locus in strains

TABLE 2. *pabA*-dependent growth rate assay

Strain	*kan* insertion[a]	Promoter(s)[b]		*pabA*[c]	Doubling time[d]	
		P_1	P_2		Without PABA	With PABA
MC1000	None	+	+	Wild type	60	62
BN198	*Bgl*II	−	+	*pabA::lacZ*	NG[e]	61
BN501	*Bgl*II	+	+	Wild type	58	60
BN502	*Bgl*II	−	+	Wild type	180	63
BN503	*Pvu*II$_b$	+	−	Wild type	67	65

[b] See Fig. 5.
[b] *pabA* transcription by the indicated promoters.
[c] Genotype denotes coding region.
[d] In minutes; PABA was supplemented at 10 ng/ml in Vogel-Bonner minimal medium.
[e] NG, No growth.

BN501 and BN502 was further analyzed by Southern analysis. The results indicated no gross structural changes at the *pabA* locus other than the addition of the *Bgl*II::*kan* marker (data not shown). In addition, the deletion of the −10 region of P_1 in strain BN502 was confirmed by sequencing analysis of the chromosomal region between the *Pvu*II$_b$ and *Hin*cII sites (Fig. 1A) amplified by the polymerase chain reaction technique (data not shown).

Growth rate analysis in minimal medium indicated that strain BN502 (ΔP_1 P_2^+) grew three times more slowly than strains BN501 (P_1^+ P_2^+) and BN503 (P_1^+ P_2^-) when PABA was omitted from the medium (Table 2). PABA synthesis in strain BN502 is due to PabA expressed from P_2, and not to possible residual PABA present in the media or to the ammonia-dependent (*pabA*-independent) system catalyzed by component I of PABS (encoded by *pabB* [14, 39]), since the isogenic null *pabA* mutant (strain BN198) did not grow at all in the same analysis (Table 2).

The result therefore indicated that PabA is not efficiently expressed from the P_2-initiated transcript, which is in agreement with the β-galactosidase assay. This level of expression was not sufficient for cellular PABA synthesis in the absence of P_1, since the decreased growth rate reflected PABA starvation.

DISCUSSION

This work investigated the chromosomal organization at the *pabA* locus in both *E. coli* and *S. typhimurium* and *pabA* expression in *E. coli*. Nucleotide sequence and transcription analyses indicated that *pabA* is encoded in two overlapping transcriptional units: the P_2-initiated polycistronic unit which includes *orf1-fic-pabA* and the P_1-initiated monocistronic unit which includes only *pabA*. P_2 lies in the intergenic region between an upstream transcriptional unit (*utu*) and *orf1* (Fig. 1 and 2). Upstream of P_2 lies a putative ρ-independent transcription terminator which appears to effect transcription termination of *utu* within the intergenic region. P_1 lies in the coding region of *fic*, with the deduced −10 region overlapping the stop codon of *fic* and the −35 region overlapping *fic* codons 193 and 194 (Fig. 2). Both promoters contained good −10 but poor −35 sequences in comparison with the *E. coli* promoter consensus sequences (13). The relative transcriptional activities of P_{utu}, P_2, P_1, and P_{dtu}, assayed by a *galK* fusion expression system (Fig. 3), were 10:1:1:3. Although transcription termination sites of the two *pabA* units have not been determined in this study, it is likely that they map in the intergenic region between *pabA* and a downstream transcription unit (*dtu*) since this region is 85 bp long and AT rich (Fig. 1 and 2).

Expression of PabA by the two transcriptional units was investigated by using directed mutagenesis and gene fusion approaches. Specifically, assessing PabA expression by P_1 alone was possible by blocking transcription from P_2 into *pabA* with a polar *kan* cassette insertion in *fic*. Expression by P_2 alone was effected by deleting the −10 sequence of P_1 with oligonucleotide-directed mutagenesis. PabA expression from either promoter was determined by using two independent but complementary assays. The first utilized the chromosomal *pabA-lacZ* protein fusion system to determine the relative level of PabA synthesis deduced from β-galactosidase activity (37). The protein fusion was constructed at the chromosomal *pabA* locus to obtain genuine PabA expression and to avoid any possible complication that may be generated when plasmid or lysogenic phage expression systems are used. The second approach measured cellular growth rate, which reflects the level of PABA synthesized by PABS.

Both assays indicated that PabA is primarily expressed from P_1 and that P_2 is dispensable for cellular PABA metabolism. Specifically, deletion of P_2 had no effect on PabA expression for cell growth, whereas in the absence of P_1, growth was reduced as cells are starving for PABA (Table 2; compare BN502 with BN503). Since P_2 expressed PabA at less than 2% of that expressed by P_1 and yet allowed some growth, it suggested that PabA or PABS component II is required at a very low level and that P_1 expressed an excess amount of component II relative to P_2. These data are consistent with preliminary findings suggesting that PABS component I (encoded by *pabB*) is the limiting component in PABA synthesis in vivo (G. G. Guay and B. P. Nichols, unpublished). The excess pool of PABA may be important for a cell to keep up with fast growth conditions when it is quickly dividing.

Since the relative steady-state levels of the two *pabA* transcripts are approximately the same, and yet only the P_1-initiated transcript was efficiently used for translation (Fig. 5 and 6), it appeared that inefficient expression of PabA by the P_2-initiated polycistronic transcript might be a result of posttranscriptional events. Nucleotide sequence and RNA analyses of this transcriptional unit suggested several possible mechanisms for posttranscriptional control. P_2 initiated a polycistronic transcript which, unlike the P_1 transcript, can form a stable secondary structure that sequesters the *pabA* ribosome-binding site, as has been previously suggested (18) (Fig. 2). Ribosome-binding sites contained within stable secondary structure are poorly recognized (10, 22, 33). PabA expression from the P_2-initiated transcript may also be coupled to *orf1* and *fic*, which appeared to be poorly translated on the basis of nucleotide sequence analysis. Translation initiation of *orf1* occurs at a GTG initiation codon, which is less efficient than an ATG initiation codon (23, 31). The overlap between *orf1* and *fic* may result in translational coupling, a mechanism that ensures that a promoter-distal gene is not efficiently translated unless the proximal gene is first completed (2, 30). Thus, low *fic* expression may be attained by coupling to an inefficiently translated *orf1*. Furthermore, in both *E. coli* and *S. typhimurium*, both *orf1* and *fic* contain a high incidence of codons rarely used in the translation of efficiently expressed genes (16, 34). Particularly notable is the frequency of arginine CGR codons, which appear at five times the expected frequency. It is possible that the presence of these rare codons requiring the use of minor charged tRNA species reduces expression of *orf1* and *fic* by modulating the rate of translation elongation (16, 40).

Inefficient expression of PabA by the P_2 polycistronic

transcript may be due to poor translation of the upstream genes and/or poor translation initiation of *pabA*. Inefficient translation could result in RNA degradation, i.e., functional inactivation of the message (26); this appeared to occur in the P_2 expression system. RNA primer extension data indeed indicated that the cDNA products of the P_2 transcript are heterogeneous in size (Fig. 6B), which suggested that the transcript was degraded at discrete sites. It is unlikely that these signals were due to premature termination by reverse transcriptase, since RNase protection analysis also indicated this (Fig. 6A). That is, besides generating two major protections of the *PstI-HincII* probe to generate signals of about 183 and 127 nucleotides, RNA from strains with *pabA* transcribed by P_2 (*BglII::kan* and *NruI::kan*) or by P_{utu} (ΔPvuII$_a$-PvuII$_b$) also generated other protections that corresponded to the degraded products of the P_2 transcript. This pattern was also observed when the *PvuII$_b$-PstI* probe (*fic*-coding region) was analyzed with strains transcribing *fic* (data not shown). The P_2 transcript thus appeared to be degraded at discrete sites. Experiments are in progress to further investigate this observation.

One of the motivating factors for this work was that of the nine biochemical reactions requiring the products of at least 10 genes to synthesize dihydrofolate from GTP, chorismate, and glutamine (9), only four genes have been identified and mapped to date. *pabA* (74 min) and *pabB* (40 min) encode the two subunits of PABS (15, 43), *folA* (1 min) encodes dihydrofolate reductase (8, 35), and *folC* (56 min) encodes the bifunctional folylpoly-γ-glutamate synthetase-dihydrofolate synthetase (6). The unlinked nature of the genes characterized so far, including those belonging to the same enzyme complex, leaves open the possibility that *pab* and *fol* genes lie in complex operons along with other genes of different functions.

Such a case was demonstrated in this study, which showed that *pabA*, besides being independently expressed, was coexpressed at a reduced level in a complex operon including two other genes, *orf1* and *fic*, of unknown identity (although *fic* is thought to encode a gene product involved in cell division [20]). Analysis of complex operons such as the macromolecular synthesis (38) and the *hisT* (4) operons has suggested that complex organization may serve to coregulate the expression of genes encoding different functions under some specific conditions. It would therefore be of interest to investigate PabA expression from P_2 under physiological growth conditions other than those used in the present study.

ACKNOWLEDGMENTS

We thank D. Dix for the mini-Mu d experiments, M. Winkler for discussion and suggestions, and J. Green and W. Merkel for critically reading the manuscript.

This work was supported by research grants from Public Health Service grants AI18639 and AI25106 from the National Institutes of Health.

LITERATURE CITED

1. **Adams, C. W., and G. W. Hatfield.** 1984. Effects of promoter strengths and growth conditions on copy number of transcription-fusion vectors. J. Biol. Chem. **259**:7399–7403.
2. **Aksoy, S., C. L. Squires, and C. Squires.** 1984. Translational coupling of the *trpB* and *trpA* genes in the *Escherichia coli* tryptophan operon. J. Bacteriol. **157**:363–367.
3. **Arps, P. J., C. C. Marvel, B. C. Rubin, D. A. Tolan, E. E. Penhoet, and M. E. Winkler.** 1985. Structural features of the *hisT* operon of *Escherichia coli* K-12. Nucleic Acids Res. **13**:5297–5315.
4. **Arps, P. J., and M. E. Winkler.** 1987. Structure analysis of the *Escherichia coli* K-12 *hisT* operon by using a kanamycin resistance cassette. J. Bacteriol. **169**:1061–1070.
5. **Ausebel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl (ed.).** 1989. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
6. **Bognar, A. L., C. Osborne, B. Shane, S. C. Singer, and R. Ferone.** 1985. Folylpoly-γ-glutamate synthetase-dihydrofolate synthetase. Cloning and high expression of the *Escherichia coli folC* gene and purification and properties of the gene product. J. Biol. Chem. **260**:5625–5630.
7. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, 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.
8. **Breeze, A. S., P. Sims, and K. A. Stacey.** 1975. Trimethoprim-resistant mutants of *E. coli* K-12: preliminary genetic mapping. Genet. Res. **25**:207–214.
9. **Brown, G. M., and J. M. Williamson.** 1981. Biosynthesis of riboflavin, folic acid, thiamine, and pantothenic acid. Adv. Enzymol. **53**:345–381.
10. **Cone, K. C., and D. A. Steege.** 1985. Functional analysis of *lac* repressor restart sites in translational initiation and reinitiation. J. Mol. Biol. **186**:733–742.
11. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. **Goncharoff, P., and B. P. Nichols.** 1984. Nucleotide sequence of *pabB* indicates common evolutionary origin of *para*-aminobenzoate synthase and anthranilate synthase. J. Bacteriol. **159**: 57–62.
13. **Hawley, D. K., and W. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
14. **Huang, M., and F. Gibson.** 1970. Biosynthesis of 4-aminobenzoate in *Escherichia coli*. J. Bacteriol. **102**:767–773.
15. **Huang, M., and J. Pittard.** 1967. Genetic analysis of mutant strains of *Escherichia coli* requiring *para*-aminobenzoic acid for growth. J. Bacteriol. **93**:1938–1942.
16. **Ikemura, T.** 1985. Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. **2**:13–34.
17. **Jones, H. M., C. M. Brajkovich, and R. P. Gunsalus.** 1983. In vivo 5′ terminus and length of the mRNA for the proton-translocating ATPase (*unc*) operon of *Escherichia coli*. J. Bacteriol. **155**:1279–1287.
18. **Kaplan, J. B., W. K. Merkel, and B. P. Nichols.** 1985. Evolution of glutamine amidotransferase genes: nucleotide sequence of the *pabA* genes from *Salmonella typhimurium*, *Klebsiella aerogenes*, and *Serratia marcescens*. J. Mol. Biol. **183**:327–340.
19. **Kaplan, J. B., and B. P. Nichols.** 1983. Nucleotide sequence of *Escherichia coli pabA* and its evolutionary relationship to *trp(G)D*. J. Mol. Biol. **168**:451–468.
20. **Kawamukai, M., H. Matsuda, W. Fujii, T. Nishida, Y. Izumoto, M. Himeno, R. Utsumi, and T. Komano.** 1988. Cloning of the *fic-1* gene involved in cell filamentation induced by cyclic AMP and construction of a Δ*fic Escherichia coli* strain. J. Bacteriol. **170**:3864–3869.
21. **Kawamukai, M., H. Matsuda, W. Fujii, R. Utsumi, and T. Komano.** 1989. Nucleotide sequences of *fic* and *fic-1* genes involved in cell filamentation induced by cyclic AMP in *Escherichia coli*. J. Bacteriol. **171**:4525–4529.
22. **Looman, A. C., J. Bodlaender, M. de Gruyter, A. Vogelaar, and P. H. van Knippenberg.** 1986. Secondary structure as primary determinant of the efficiency of ribosomal binding sites in *Escherichia coli*. Nucleic Acids Res. **14**:5481–5497.
23. **Looman, A. C., and P. H. van Knippenberg.** 1986. Effects of GUG and AUG initiation codons on the expression of *lacZ* in *Escherichia coli*. FEBS Lett. **197**:315–320.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled

DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.

26. **McCarthy, J. E. G., B. Schauder, and P. Ziemke.** 1988. Post-transcriptional control in *Escherichia coli*: translation and degradation of the *atp* operon mRNA. Gene **72**:131–139.
27. **McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady, and M. Rosenberg.** 1981. A system to study promoter and terminator signals recognized by *Escherichia coli* RNA polymerase, p. 383–415. *In* J. G. Chirikjian and T. S. Papas (ed.), Gene amplification and analysis, vol. 2. Elsevier/North-Holland Publishing Co., New York.
28. **Messing, J., and J. Vieira.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:269–276.
29. **Miller, J.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

29a. **Nichols, B. R., A. M. Seibold, and S. Z. Doktor.** 1989. *para*-Aminobenzoate synthesis from chorismate occurs in two steps. J. Biol. Chem. **264**:8597–8601.

30. **Oppenheim, D. S., and C. Yanofsky.** 1980. Translational coupling during expression of the tryptophan operon of *Escherichia coli*. Genetics **95**:785–795.
31. **Reddy, P., A. Peterkofsky, and K. McKenny.** 1985. Translational efficiency of the *Escherichia coli* adenylate cyclase gene: mutating the UUG initiation codon to GUG or AUG results in increased gene expression. Proc. Natl. Acad. Sci. USA **82**: 5656–5660.
32. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
33. **Schneider, K., and C. F. Beck.** 1988. Point mutations that affect translation initiation of the transposon Tn10 *tetA* gene. Gene **74**:559–563.
34. **Sharp, P. M., and W.-H. Li.** 1986. Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for "rare" codons. Nucleic Acids Res. **14**:7737–7749.
35. **Sheldon, R.** 1977. Altered dihydrofolate reductase in *fol* regulatory mutants of *Escherichia coli* K-12. Mol. Gen. Genet. **151**: 215–219.
36. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
37. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
38. **Taylor, W. E., D. B. Straus, A. D. Grossman, Z. F. Burton, C. A. Gross, and R. Burgess.** 1984. Transcription from a heat-inducible promoter causes heat shock regulation of the sigma subunit of *E. coli* RNA polymerase. Cell **38**:371–381.
39. **Teng, C.-Y. P., B. Ganem, S. Z. Doktor, B. P. Nichols, R. K. Bhatnagar, and L. C. Vining.** 1985. Total synthesis of (+/−)4-amino-4-deoxychorismic acid: a key intermediate in the biosynthesis of para-aminobenzoic acid and L-(para-aminophenyl)alanine. J. Am. Chem. Soc. **107**:5008–5009.
40. **Thomas, L. K., D. B. Dix, and R. C. Thompson.** 1988. Codon choice and gene expression: synonymous codons differ in their ability to direct aminoacylated-transfer RNA binding to ribosomes *in vitro*. Proc. Natl. Acad. Sci. USA **85**:4242–4246.
41. **Vogel, H. J., and D. M. Bonner.** 1956. Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. **218**:97–106.
42. **Weinstock, G. M., M. L. Berman, and T. J. Silhavy.** 1983. Chimeric genetics with β-galactosidase, p. 27–64. *In* T. S. Papas, M. Rosenberg, and J. G. Chirikjian (ed.), Gene amplification and analysis. Elsevier/North-Holland Publishing Co., New York.
43. **White, B. J., S. J. Hochhauser, N. M. Cintron, and B. Weiss.** 1976. Genetic mapping of *xthA*, the structural gene for exonuclease III in *Escherichia coli*. J. Bacteriol. **126**:1082–1088.
44. **Winans, S. C., S. J. Elledge, J. H. Krueger, and G. C. Walker.** 1985. Site-directed insertion and deletion mutagenesis with cloned fragments in *Escherichia coli*. J. Bacteriol. **161**:1219–1221.