JOURNAL OF BACTERIOLOGY, Sept. 1989, p. 4767–4777
0021-9193/89/094767-11$02.00/0

Vol. 171, No. 9

Overlap between *pdxA* and *ksgA* in the Complex *pdxA-ksgA-apaG-apaH* Operon of *Escherichia coli* K-12

BENJAMIN B. ROA, DENNIS M. CONNOLLY, AND MALCOLM E. WINKLER*

Department of Molecular Biology, Northwestern University Medical School, Chicago, Illinois 60611

Received 23 February 1989/Accepted 9 June 1989

We report that *pdxA*, which is required for de novo biosynthesis of pyridoxine (vitamin B_6) and pyridoxal phosphate, belongs to an unusual, multifunctional operon. The *pdxA* gene was cloned in the same 3.5-kilobase *Bam*HI-*Eco*RI restriction fragment that contains *ksgA*, which encodes the 16S rRNA modification enzyme m_2^6A methyltransferase, and *apaH*, which encodes diadenosine tetraphosphatase (AppppA hydrolase). Previously, Blanchin-Roland et al. showed that *ksgA* and *apaH* form a complex operon (Mol. Gen. Genet. 205:515–522, 1986). The *pdxA* gene was located on recombinant plasmids by subcloning, complementation, and insertion mutagenesis, and chromosomal insertions at five positions upstream from *ksgA* inactivated *pdxA* function. DNA sequence analysis and minicell translation experiments demonstrated that *pdxA* encoded a 35.1-kilodalton polypeptide and that the stop codon of *pdxA* overlapped the start codon of *ksgA* by 2 nucleotides. The translational start codon of *pdxA* was tentatively assigned based on polypeptide size and on the presence of a unique sequence that was also found near the translational start of PdxB. This conserved sequence may play a role in translational control of certain pyridoxine biosynthetic genes. RNase T_2 mapping of chromosomal transcripts confirmed that *pdxA* and *ksgA* were members of the same complex operon, yet about half of *ksgA* transcripts arose in vivo under some culture conditions from an internal promoter mapped near the end of *pdxA*. Transcript analysis further suggested that *pdxA* is not the first gene in the operon. These structural features support the idea that pyridoxine-biosynthetic genes are members of complex operons, perhaps to interweave coenzyme biosynthesis genetically with other metabolic processes. The results are also considered in terms of *ksgA* expression.

Pyridoxal phosphate is used as a coenzyme by dozens of enzymes that participate in all phases of amino acid metabolism (6). These enzymes have widely varying affinities for this essential coenzyme, which suggests that pyridoxal phosphate availability may modulate enzyme activities and thereby control the flow of intermediates down metabolic pathways. In addition, pyridoxal phosphate may function as an allosteric effector of metabolic pathways that do not directly involve amino acids (23, 42). Because of these diverse metabolic roles and its intrinsic reactivity with numerous proteins (19), it is likely that cellular levels of pyridoxal phosphate are carefully controlled (15, 17). In certain bacteria, fungi, and plants, the pyridoxal phosphate precursor pyridoxine (vitamin B_6) is synthesized de novo from common metabolic intermediates (25, 39, 40). According to one scheme, pyridoxine is then phosphorylated to pyridoxine phosphate and oxidized to pyridoxal phosphate by a series of steps found in all organisms (18). Because pyridoxine biosynthesis involves precursors that are shared by several other pathways, determination of the exact pyridoxine-biosynthetic pathway has been hampered despite rigorous analysis by isotopic labeling experiments (25).

So far, only two pyridoxine-biosynthetic (*pdx*) genes have been studied in detail. Interestingly, both *pdxB* and *serC* (*pdxF*) are members of complex operons in that they form transcriptional units with genes whose functions are not directly linked to coenzyme biosynthesis. The *pdxB* gene is the first gene in a five-member operon that includes *hisT*, which encodes the modification enzyme tRNA pseudouridine synthase I (4, 28; P. Schoenlein, B. Roa, and M. Winkler, submitted for publication). The *pdxB-hisT* operon also contains a gene that shows striking homology to the *Streptococcus mutans* gene encoding aspartate semialdehyde dehydrogenase (24), a gene that encodes an integral membrane protein, and a gene that encodes a membrane-associated, zinc finger-containing regulatory protein (10, 33) (D. Connolly and M. Winkler, unpublished observations). The *serC*(*pdxF*) gene product is bifunctional and is required for both serine and pyridoxine biosynthesis (20). The *serC* (*pdxF*) gene forms a complex operon with *aroA*, which encodes an important enzyme in prechorismate biosynthesis (21).

For all complex operons, a basic assumption has been that genes with apparently different functions are organized together into overlapping transcriptional and translational units to allow coregulation under some physiological conditions and thereby ensure coordination of cellular metabolism (7). Based on the unusual genetic link between *pdxB* and *hisT*, we proposed that *pdx* biosynthetic genes may be grouped into complex operons to interweave coenzyme biosynthesis with other metabolic processes (3). To test this notion and to learn more about the regulation of pyridoxal phosphate biosynthesis, we determined the structure of *pdxA*, which, like *pdxB*, plays a role in pyridoxine biosynthesis. In this paper, we show that *pdxA* forms a complex operon with *ksgA*, which encodes the modification enzyme 16S rRNA m_2^6A methyltransferase (43). Mutants defective in *ksgA* function lack methyl groups on two adenosine residues near the 3′ end of 16S rRNA and are therefore resistant to the antibiotic kasugamycin. Previously, Fayat and coworkers demonstrated that *ksgA* itself is in a complex operon with *apaG* and *apaH* (8). The *apaH* gene encodes diadenosine tetraphosphatase, which catalyzes the breakdown of the important alarmone AppppA (29). Thus, the results presented here establish genetic links between pyridoxal phosphate biosynthesis and several far-reaching phys-

* Corresponding author.

iological processes. In addition, we report features of *pdxA* structure and in vivo transcription that are involved in expression of the *pdxA-ksgA-apaG-apaH* operon and that might play roles in translational coordination of enzymes involved in pyridoxine biosynthesis.

MATERIALS AND METHODS

Materials. An *Escherichia coli* K-12 genomic library containing *Bam*HI inserts cloned into pBR322 was a gift from B. Nichols (University of Illinois, Chicago). Restriction enzymes, T4 DNA ligase, T4 DNA polymerase, and M13mp18 and M13mp19 phage vectors were purchased from New England BioLabs (Beverly, Mass.). Plasmid vectors pUC18 and pUC19, deoxynucleotide triphosphates, dideoxynucleotide triphosphates, and single-stranded 17-mer sequencing primer were bought from Pharmacia Inc. (Piscataway, N.J.). Culture media were from Difco Laboratories (Detroit, Mich.). Amino acids, antibiotics, biochemical reagents, and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) protein molecular weight standards were purchased from Sigma Chemical Co. (St. Louis, Mo.). L-[^{35}S]methionine (≈1,260 Ci/mmol), [α-^{32}P]dCTP (>800 Ci/mmol), [α-^{32}P]dATP (>3,000 Ci/mmol), and [α-^{32}P]CTP (>400 Ci/mmol) were purchased from Amersham Corp. (Arlington Heights, Ill.). The deoxynucleotide triphosphate analog 7-deaza-dGTP was obtained from Boehringer Mannheim Biochemicals (Indianapolis, Ind.). DNA polymerase I large (Klenow) fragment and "Sequenase" kits were purchased from United States Biochemical Corp. (Cleveland, Ohio). Reagents used in preparing sequencing gels included ammonium persulfate and acrylamide (Bio-Rad Laboratories, Richmond, Calif.), formamide (Fisher Scientific, Fair Lawn, N.J.), and urea (Sigma). The Riboprobe Gemini system, pGEM-3Z cloning vector, and RQ1 DNase (RNase-free) were purchased from Promega Biotec (Madison, Wis.). RNase T_2 was from Bethesda Research Laboratories, Gaithersburg, Md.

Bacterial strains, media, and culture conditions. Bacterial strains are listed in Table 1. Markers were moved between strains by generalized transduction with P1 *kc* phage (31). Kanamycin resistance cassettes (Km^R) cloned into specific restriction sites and mini-Mu dl(1734-Km^r) transcriptional fusions were crossed from linearized plasmids into the chromosome of *recBC sbc* mutant JC7623 by the method of Winans and Walker, as described previously (45). Km^R insertion mutations or mini-Mu dl(1734-Km^r) fusions were transduced into a W3110 prototroph, NU426, or a W3110 *tnaA2* Δ*lacU169* mutant, NU816, respectively.

Bacteria were cultured in LB medium supplemented with 30 μg of cysteine per ml (LB+Cys) and in Vogel-Bonner minimal (E) medium (13) supplemented with 0.4% glucose and 10^{-5} M $FeSO_4$ (minimal E-Glc-Fe). Standard LB+Cys medium contains sufficient concentrations of pyridoxine and pyridoxal to support full growth of pyridoxine-requiring auxotrophs. When necessary, pyridoxine or pyridoxal was added to minimal E-Glc-Fe medium at the "physiological" concentration of 5×10^{-7} M (15). Other supplements were added at concentrations suggested elsewhere (13, 31, 41). If required, ampicillin and kanamycin were added to LB+Cys or minimal E-Glc-Fe medium at 50 and 25 μg/ml, respectively. Kasugamycin was added as indicated in Table 3. Culture temperatures are listed in the tables and figures. Compounds were tested for their abilities to support growth of *pdxA* and *pdxB* mutants as detailed in the footnotes to Table 2.

Plasmids and M13 phage clones. Table 1 also lists recombinant plasmids used in this study. Restriction fragment isolations, DNA manipulations, and cloning were carried out by standard protocols (27). Restriction fragments with noncompatible ends were filled in with Klenow enzyme (5′ overhang) or chewed back with T4 DNA polymerase (3′ overhang) to yield blunt-ended fragments suitable for cloning (27). The kanamycin resistance cassette (Km^R), which was cloned into the single *Eco*RV site of plasmids pNU121, pNU122, and pNU123 to give plasmids pNU125 through pNU132 (Table 1), was from plasmid pMB2190, and its properties were described earlier (4). The direction of transcription of the inserted *kan* gene was determined in each construct by restriction analysis (4, 27). Mini-Mu dl(1734-Km^r) elements were jumped into plasmid pNU167 as described before (11). Four independent transcriptional fusions (plasmids pNU209 through pNU212) that inactivated *pdxA* and showed *lacZ* expression on MacConkey-lactose medium were mapped with restriction enzymes (27). Plasmids pNU180, pNU181, and pNU182 were constructed by cloning into plasmid pGEM-3Z and used as templates for RNA probes. DNA fragments indicated in Fig. 1 were cloned into bacteriophage vectors M13mp18 and M13mp19 essentially as described before (30, 46).

Minicell translation experiments. Plasmid-encoded gene products were labeled with L-[^{35}S]methionine in purified minicells, resolved on 12.5% linear or 5 to 20% gradient polyacrylamide gels (22), and detected by autoradiography as described previously (28). Molecular mass standards included β-galactosidase (116 kilodaltons [kDa]), fructose-6-phosphate kinase (84 kDa), pyruvate kinase (58 kDa), fumarase (48.5 kDa), lactate dehydrogenase (36.5 kDa), triose phosphate isomerase (26.6 kDa), and lysozyme (14.3 kDa).

DNA sequence determinations. DNA sequences were determined independently by two related methods. In both cases, conditions were optimized to prevent elongation blocks and gel compressions. The first method contained the following modifications of the Sanger chain termination method (36): (i) [α-^{32}P]dCTP replaced [α-^{32}P]dATP, with appropriate changes in compositions of reaction mixtures; (ii) elongation and chase reaction mixes were incubated for 10 min at 37°C; and (iii) 7-deaza-dGTP replaced dGTP in elongation reaction mixtures. The second method was a variation of the Sequenase kit reactions provided by the manufacturer and included deoxyinosine triphosphate (dITP) and single-strand-binding protein (SSB) in the reaction mixes. DNA products from sequencing reactions were resolved on 80-cm 6% polyacrylamide gels containing TBE buffer, 8 M urea, and 20% (vol/vol) formamide (35). DNA sequence analyses were performed with UWGCG (University of Wisconsin Genetics Computer Group), PCGene (IntelliGenetics, Inc.), and BIONET computer programs.

RNase T_2 transcript analysis. Strain NU426 or NU812 was grown in LB+Cys or minimal E-Glc-Fe medium containing 5×10^{-7} M pyridoxine, respectively, at 37°C with shaking. Total cellular RNA was isolated from each strain by a modification of the SDS-hot phenol method and treated with RQ1 DNase as described before (4). RNA probes (see Fig. 1) were synthesized with the Riboprobe Gemini system according to the manufacturer's instructions with [α-^{32}P]CTP as the labeled nucleoside triphosphate. RNase T_2 mapping of in vivo transcripts was performed as described before (12) with the following modifications: (i) 50 μg of RQ1 DNase-treated RNA and 250,000 cpm (Cerenkov) of ^{32}P-labeled RNA probe were used per hybridization reaction; (ii) hybridization re-

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Genotype[a]	Source or reference
E. coli K-12		
CGSC4541 (AT2365)	*ara galK lacY leu mtl-1 pro pdxA4*::Mu *thr thi strA20 supE44 xyl*	A. L. Taylor, in B. Bachmann collection
CGSC6602 (FS131)	*gal ksgA19 lacY metB mtl-2 purF1 rimG21 supE44 tonA2 tsx-1 xyl*	P. F. Sparling, in B. Bachmann collection
JC7623	*arg ara his leu pro recB21 recC22 sbcB15 thr*	A. J. Clark (45)
JM105	*thi rpsL* Δ(*lac-proAB*) *endA sbc-15 hspR4* (F′ *traD36 proAB lacI*^{q} *Z*Δ*M15*)	Pharmacia LKB (46)
JM109	*recA1 endA1 gyrA96 thi hsdR17 supE44 relA1* Δ(*lac-proAB*) (F′ *traD36 proAB lacI*^{q} *Z*Δ*M15*)	Promega Biotec (46)
P678-54	*ara azi gal leu lacY malA minA minB rpsL thi thr tonA tsx xyl*	P. Matsumura (5)
NU426	W3110 prototroph	C. Yanofsky collection
NU687	JC7623 *pdxA*::Km^R(*Eco*RV)>	Transformation with linearized pNU128
NU688	JC7623 *pdxA*::<Km^R(*Eco*RV)	Transformation with linearized pNU129
NU689	CGSC4541 *srl*::Tn*10 recA1*	NU638 × P1 *kc* (TT9813)
NU690	CGSC6602 *srl*::Tn*10 recA1*	NU640 × P1 *kc* (TT9813)
NU811	NU426 *pdxA*::Km^R(*Eco*RV)>	NU426 × P1 *kc* (NU687)
NU812	NU426 *pdxA*::<Km^R(*Eco*RV)	NU426 × P1 *kc* (NU688)
NU816	Single-colony isolate of W3110 *tnaA2* Δ*lacU169*	C. Yanofsky collection
NU1126	JC7623 *pdxA*::mini-Mu d1-1 Km^r	Transformation with linearized pNU209
NU1127	JC7623 *pdxA*::mini-Mu d1-2 Km^r	Transformation with linearized pNU210
NU1187	NU816 *pdxA*::mini-Mu d1-1 Km^r	NU816 × P1 *kc* (NU1126)
NU1188	NU816 *pdxA*::mini-Mu d1-2 Km^r	NU816 × P1 *kc* (NU1127)
TT9813	*ara eda* Δ*lacU169 metF*(Am) *mtl-1 recA1 rpsL srl*::Tn*10 thi tonA31 tsx xyl*	Hughes and Roth (26)
Plasmids		
pACYC184	Replicon P15A; Cm^r Tc^r	Chang and Cohen (27)
pBR322	Replicon ColE1; Ap^r Tc^r	Bolivar et al. (27)
pGEM3Z	pUC18 derivative with phage T7 and Sp6 promoters	Promega Biotec
pMB2190	Km^R in pBR327 derivative; Ap^r Tc^r	B. Nichols collection
pUC18	Replicon ColE1; LacZ$^+$ Ap^r	Pharmacia LKB (46)
pUC19	Same as pUC18, but with opposite orientation of polylinker	Pharmacia LKB (46)
pNU121	pUC18 [*Bam*HI(1)-*Eco*RI(3477)] Ap^r	Subclone of pNU244
pNU122	pUC19 [*Bam*HI(1)-*Eco*RI(3477)] Ap^r	Subclone of pNU244
pNU123	pBR322 [*Bam*HI(1)-*Eco*RI(3477)] Ap^r	Subclone of pNU244
pNU125	pNU122 [*pdxA*::<Km^R(*Eco*RV)]; Ap^r Km^r	Subclone with Km^R cassette in *Eco*RV(810) site
pNU128	pNU123 [*pdxA*::Km^R(*Eco*RV)>]; Ap^r Km^r	Subclone with Km^R cassette in *Eco*RV(810) site
pNU129	pNU123 [*pdxA*::<Km^R(*Eco*RV)]; Ap^r Km^r	Subclone with Km^R cassette in *Eco*RV(810) site
pNU130	pNU121 [*pdxA*::Km^R(*Eco*RV)>]; Ap^r Km^r	Subclone with Km^R cassette in *Eco*RV(810) site
pNU131	pNU121 [*pdxA*::<Km^R(*Eco*RV)]; Ap^r Km^r	Subclone with Km^R cassette in *Eco*RV(810) site
pNU132	pNU122 [*pdxA*::Km^R(*Eco*RV)>]; Ap^r Km^r	Subclone with Km^R cassette in *Eco*RV(810) site
pNU167	pACYC184 [*Bam*HI(1)-*Hpa*I(1275)]; Cm^r	Subclone into *Bam*HI-*Hin*cII-cut vector
pNU180	pGEM3Z [*Bgl*I(1081)-*Pvu*II(1466)]; Ap^r LacZ$^-$	See Results
pNU181	pGEM3Z [*Eco*RV(810)-*Bgl*I(1081)]; Ap^r LacZ$^-$	See Results
pNU182	pGEM3Z [*Bam*HI(1)-*Ava*II(655)]; Ap^r LacZ$^-$	See Results
pNU209	pNU167 *pdxA*::mini-Mu d1-1 Cm^r Km^r	Mini-Mu d1 jump into pNU167 (11)
pNU210	pNU167 *pdxA*::mini-Mu d1-2 Cm^r Km^r	Mini-Mu d1 jump into pNU167 (11)
pNU211	pNU167 *pdxA*::mini-Mu d1-3 Cm^r Km^r	Mini-Mu d1 jump into pNU167 (11)
pNU212	pNU167 *pdxA*::mini-Mu d1-4 Cm^r Km^r	Mini-Mu d1 jump into pNU167 (11)
pNU243	pBR322 [*Bam*HI(1)-*Pvu*II(1469)] Ap^r	Subclone of pNU121
pNU244	*pdxA*$^+$ in pBR322; Ap^r	Clone from *Bam*HI genomic library

[a] < or > indicates that the direction of transcription of *kan* in the Km^R cassette is opposite to or the same as that of *pdxA*, respectively. Restriction sites refer to the map in Fig. 1, where *Bam*HI(1) and *Eco*RI(3477) are at left and right ends, respectively. Km^R(restriction site) signifies the presence of a kanamycin resistance cassette cloned into that site, whereas Km^r used as a phenotype indicates kanamycin resistance. Ap^r, Ampicillin resistant; Cm^r, chloramphenicol resistant; Tc^r, tetracyline resistant. Mini-Mu d1 is mini-Mu d1(1734-Km^r) from reference 11.

action mixes were incubated for 12 to 16 h at 50°C instead of at 60°C; and (iii) 300 μl of 120-U/ml instead of 60-U/ml RNase T_2 was added to each hybridization mixture. RNA transcripts were resolved on 40-cm 6% polyacrylamide gels containing TBE, 8 M urea, and 20% formamide, as described above for DNA sequencing.

RESULTS

Genetic localization of *pdxA*. Previously, Andresson and Davies showed that *pdxA* and *ksgA* were located in a 6.1-kilobase (kb) *Eco*RI fragment from the region around 1 min in the *E. coli* chromosome (1). Blanchin-Roland et al. used DNA sequence analysis to locate *ksgA*, *apaG*, *apaH*, and the end of *folA* to the 3.5-kb *Bam*HI-*Eco*RI fragment shown in Fig. 1 (8). To clone *pdxA*, we transformed *pdxA4*::Mu$^+$ mutant CGSC4541 with a *Bam*HI *E. coli* genomic library constructed in plasmid pBR322. A recombinant plasmid, designated pNU244, was selected that complemented the pyridoxine requirement of CGSC4541. Restriction analysis of pNU244 revealed a 13-kb chromosomal insert that

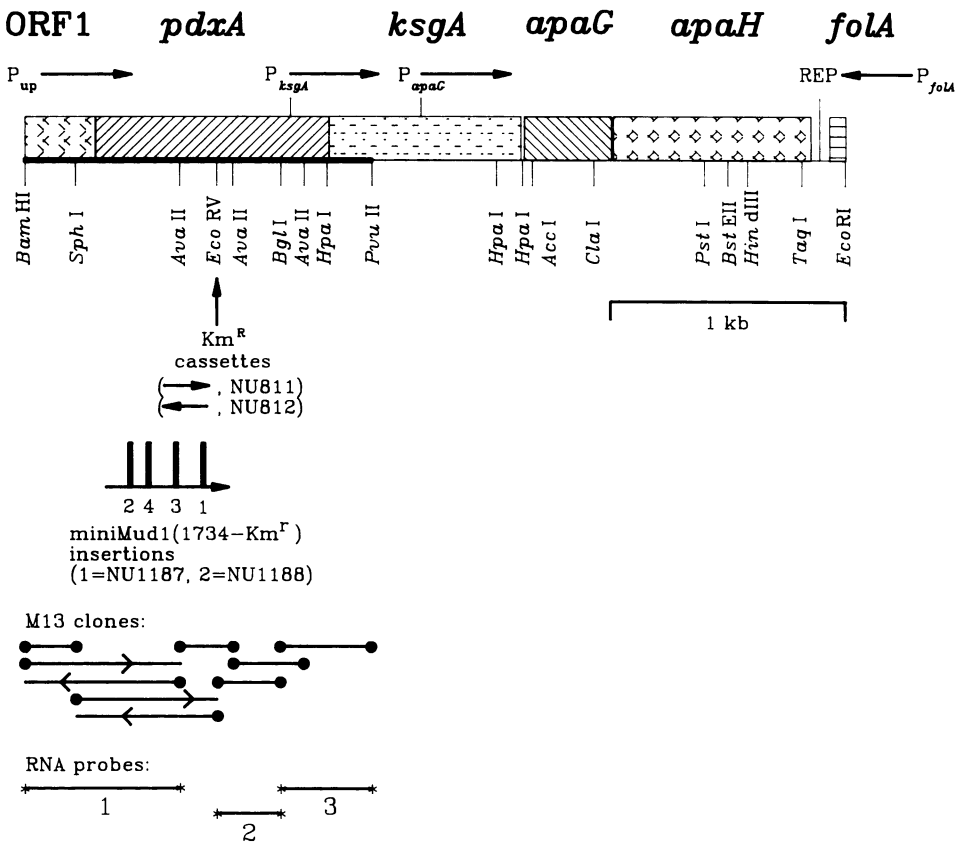

FIG. 1. Structure of the complex *pdxA-ksgA-apaG-apaH* operon of *Escherichia coli* K-12. The region analyzed in this paper, which contains ORF1 and *pdxA*, is indicated by the heavy black line. The remainder of the operon structure was taken from Blanchin-Roland et al. (8). The positions of internal promoters P_{ksgA} and P_{apaG} are indicated, based on data presented here and elsewhere (8). The P_{up} promoter must be located upstream of the *Bam*HI site (see Results). The vertical arrow highlights the position of chromosomal *pdxA*::Km^R insertion mutations in strains NU811 and NU812 (Table 1); the arrow represents the direction of *kan* gene transcription. Numbered rectangles on the horizontal arrow mark positions of four mini-Mu dl(1734-Km^r) insertions that formed transcriptional fusions in plasmids pNU209 through pNU212 (Table 1). The direction of the arrow indicates the orientation of *lacZY* that gave a Lac^+ phenotype in Δ*lacU169* mutants. Fusions 1 and 2 were crossed into the bacterial chromosome in strains NU1187 and NU1188 (Table 1). In the DNA sequencing strategy, fragments that were determined completely on both strands are indicated by lines with solid circles at both ends. On the remaining "overlap" fragments, arrowheads mark the extents to which sequences were read on one strand. The *Bam*HI-*Ava*II, *Eco*RV-*Bgl*I, and *Bgl*I-*Pvu*II fragments are shown, which were cloned into plasmid pGEM3Z and transcribed into RNA probes 1, 2, and 3, respectively, as described in Materials and Methods.

extended rightward from the *Bam*HI site shown at the left of Fig. 1. Because the restriction map constructed by Andresson and Davies showed that *pdxA* and *ksgA* lie to the left of the *Eco*RI site shown at the right of Fig. 1, our complementation data meant that the same *Bam*HI-*Eco*RI fragment that contains *ksgA*, *apaG*, and *apaH* must also contain *pdxA*. Subcloning this *Bam*HI-*Eco*RI fragment into pUC18, pUC19, and pBR322 to give pNU121, pNU122, and pNU123, respectively (Table 1), confirmed the presence of *pdxA*$^+$, since all three plasmids complemented the *pdxA4*::Mu^+ mutation in strain CGSC4541 and its *recA1* derivative NU689. In addition, when pNU121, pNU122, or pNU123 was transformed into the *ksgA19 recA1* mutant NU690, the resulting strains lost resistance to kasugamycin and showed decreased viability or lack of growth in the presence of 100 or 200 μg of kasugamycin per ml, respectively, in LB+Cys medium (data not shown). This level of kasugamycin sensitivity conferred by these multicopy *pdxA*$^+$ *ksgA*$^+$ plasmids was the same as for *ksgA*$^+$ strain NU426 (data not shown).

We located *pdxA* by subcloning the *Bam*HI-*Pvu*II DNA fragment, indicated by the thick line at the left of Fig. 1, into pBR322. The resulting plasmid, designated pNU243, complemented *pdxA* mutations (data not shown), which confirmed that the *ksgA*$^+$, *apaG*$^+$, and *apaH*$^+$ genes present on plasmids pNU121, pNU122, and pNU123 were not responsible for the *pdxA* complementation patterns described above. Likewise, a *Bam*HI-*Hpa*I fragment, which is 191 base pairs (bp) smaller than the *Bam*HI-*Pvu*II fragment in pNU243 (Fig. 1), complemented *pdxA* function when it was cloned into pACYC184 digested with *Bam*HI and *Hin*cII (plasmid pNU167, Table 1). Thus, *pdxA* is located in the minimal pNU167 clone within 1,275 bp of the *Bam*HI site shown at the left of Fig. 1.

We localized *pdxA* further by cloning a kanamycin resistance cassette (Km^R) in both orientations into the *Eco*RV site upstream of *ksgA* in plasmids pNU121, pNU122, and pNU123 (Fig. 1; Table 1). Plasmids containing Km^R(*Eco*RV) insertion mutations failed to complement the *pdxA4*::Mu^+ mutation in CGSG4541 or its *recA1* derivative NU689 (data not shown). Blanchin-Roland et al. determined the functions of the genes downstream from the *Bgl*I site near the Km^R (*Eco*RV) insertions (Fig. 1), and none seemed to be involved in pyridoxine biosynthesis (8). Therefore, the Km^R(*Eco*RV) insertions were most likely causing loss of *pdxA* function by

TABLE 2. Compounds that support the growth of *pdxA*::Km^R and *pdxB*::Km^R mutants[a]

Strain	Temp (°C)	Growth[b]						
		No addition	POX	PAL	PAM	GLYC	PYR	D-Ala
NU811 [*pdxA*::(*Eco*RV)>] Km^R	30	−	++	++	+	−	−	−
	37	−	++	++	+	−	−	−
	42	−	++	++	+	−	−	−
NU812 [*pdxA*::<Km^R(*Eco*RV)]	30	−	++	++	+	−	−	−
	37	−	++	++	+	−	−	−
	42	−	++	++	+	−	−	−
NU608 [*pdxB*::<Km^R(*Eco*RV)]	30	+	++	++	+	++	+	+
	37	−	++	++	+	+	−	+
	42	−	++	++	+	−	−	±

[a] Bacteria were grown overnight in LB+Cys medium at 37°C, centrifuged, washed twice in E medium salts, and spread onto prewarmed plates containing minimal E plus 0.4% glucose medium and the supplements listed below. Cells were also spread onto minimal E–0.4% glucose plates, and crystals of test compounds were applied to the plates. Growth was scored after 24 to 48 h of incubation at the indicated temperatures. Symbols: ++, dense growth; +, confluent but lighter growth; ±, marginal growth; −, no growth.

[b] Abbreviations and concentrations: POX, pyridoxine hydrochloride (5×10^{-7} M); PAL, pyridoxal hydrochloride (5×10^{-7} M); PAM, pyridoxamine dihydrochloride, 98 to 99% pure (5×10^{-7} M); GLYC, glycoaldehyde (0.3 mM); PYR, 4-pyridoxic acid (5×10^{-7} M); D-Ala, D-alanine (1 mM).

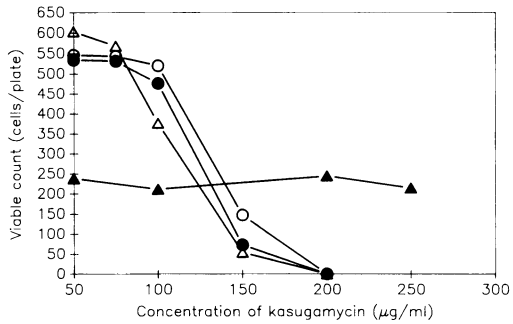

FIG. 2. Kasugamycin sensitivity of *pdxA*::Km^R mutants. Bacteria were grown overnight in LB+Cys medium at 37°C, collected by centrifugation, washed once in E medium salts, diluted in the same medium to give OD_{550} values between 0.40 and 0.55, diluted further by a factor of 10^5, and plated in 100-μl portions onto warmed LB+Cys plates containing the indicated amounts of kasugamycin. Plates were incubated at 37°C and scored after 24 to 48 h. NU426 (prototrophic parent), open triangles; NU690 (*ksgA19 recA1*) control, solid triangles; NU811 [*pdxA*::Km^R(*Eco*RV)>], open circles; NU812 [*pdxA*::<Km^R(*Eco*RV)], solid circles.

directly disrupting *pdxA* rather than through transcriptional polarity (4). Plasmids containing putative *pdxA*::Km^R (*Eco*RV) insertions still restored kasugamycin sensitivity to *ksgA19* mutants, which is consistent with the existence of a promoter immediately upstream from *ksgA* (P_{ksgA}; Fig. 1) noted previously by Blanchin-Roland et al. (8). The exact location and relative in vivo expression of the P_{ksgA} internal promoter is described below.

We also jumped mini-Mu dl(1734-Km^r) insertion elements into minimal clone pNU167 as described in Materials and Methods. Four independent transcriptional fusions (plasmids pNU209 through pNU212, Table 1) that resulted in loss of plasmid-encoded *pdxA* function are depicted in Fig. 1. These fusion plasmids were initially screened for *lacZ* expression in a Δ*lacU169* mutant by formation of red colonies on MacConkey-lactose medium. Restriction analysis showed that the mini-Mu dl(1734-Km^r) element was in the same orientation but at a different position in each plasmid (Fig. 1). This result suggested that *pdxA* transcription proceeds in the same direction as transcription of *ksgA*, *apaG*, and *apaH* (left to right in Fig. 1). DNA sequence and transcript mapping experiments presented below corroborated this direction of *pdxA* transcription.

To confirm loss of *pdxA* function, we crossed the Km^R (*Eco*RV) insertions in both orientations and two of the mini-Mu dl(1734-Km^r) transcriptional fusions into the bacterial chromosome as described in Materials and Methods. We then transduced the Km^R(*Eco*RV) chromosomal insertions or mini-Mu dl(1734-Km^r) fusions into the W3110 prototroph NU426 (to give NU811 and NU812) or into the W3110 *tnaA2* Δ*lacZU169* mutant NU816 (to give NU1187 and NU1188) (Table 1; Fig. 1). The growth characteristics of strains NU811 and NU812 are compared in Table 2 with those of strain NU608, which is a *pdxB*::Km^R mutant constructed in an earlier study (4). The data show that the Km^R(*Eco*RV) chromosomal insertions in NU811 and NU812 resulted in a nutritional requirement for the B_6 vitamers pyridoxine, pyridoxal, and pyridoxamine. Furthermore, NU811 and NU812 were restored to prototrophy by *pdxA*$^+$ *ksgA*$^+$ *apaG*$^+$ *apaH*$^+$ plasmids pNU121, pNU122, and pNU123, but not by their Km^R(*Eco*RV)-containing derivatives (data not shown). Together, these results confirm that NU811 and NU812 contain *pdxA*::Km^R(*Eco*RV) insertion mutations. Similarly, the *lacZ* transcription fusion strains NU1187 and NU1188 showed a requirement for B_6 vitamers (data not shown), which suggests that they contain *pdxA*::mini-Mu dl(1734-Km^r) insertion mutations. These *pdxA* transcriptional fusions produced about 100 Miller units of β-galactosidase activity (31) in strains NU1187 and NU1188 growing exponentially in LB+Cys medium at 37°C (data not shown). Finally, Table 2 shows that unlike the *pdxB*::Km^R mutant (3), growth of *pdxA*::Km^R(*Eco*RV) mutants was not supported by glucose at 30°C or by glucose and glycoaldehyde, 4-pyridoxic acid, or D-alanine at any temperature. The significance of this growth pattern to pyridoxine biosynthesis will be taken up in the Discussion.

Previously, we showed that the Km^R cassette used to construct *pdxA* mutations in NU811 and NU812 can cause strong transcriptional polarity when inserted into the bacterial chromosome (4). Consequently, it was of interest to determine whether NU811 and NU812 still expressed *ksgA* from the P_{ksgA} promoter, which previously had only been detected on high-copy-number recombinant plasmids (8). Figure 2 shows that the minimal inhibitory concentration of kasugamycin was essentially the same for the *pdxA*::Km^R mutants NU811 and NU812 as for their *pdxA*$^+$ *ksgA*$^+$ parent, NU426. By contrast, *ksgA19* mutant CGSC6602 grew fully at all kasugamycin concentrations tested. RNase T_2 transcript mapping experiments presented later confirmed that the Km^R cassette in the *Eco*RV site oriented with *kan* opposite to *pdxA* [*pdxA*::<Km^R(*Eco*RV) in NU812] was polar on downstream transcription. Thus, expression of *ksgA* from P_{ksgA} alone is sufficient to make cells fully sensitive to kasugamycin in LB+Cys medium.

DNA sequence of the *pdxA* region. We determined the DNA sequence of the *pdxA* region on both DNA strands by the strategy shown at the lower left of Fig. 1 and described in Materials and Methods. An analysis of chromosomal transcription presented below showed that the DNA strand shown in Fig. 3 serves as the coding strand throughout the

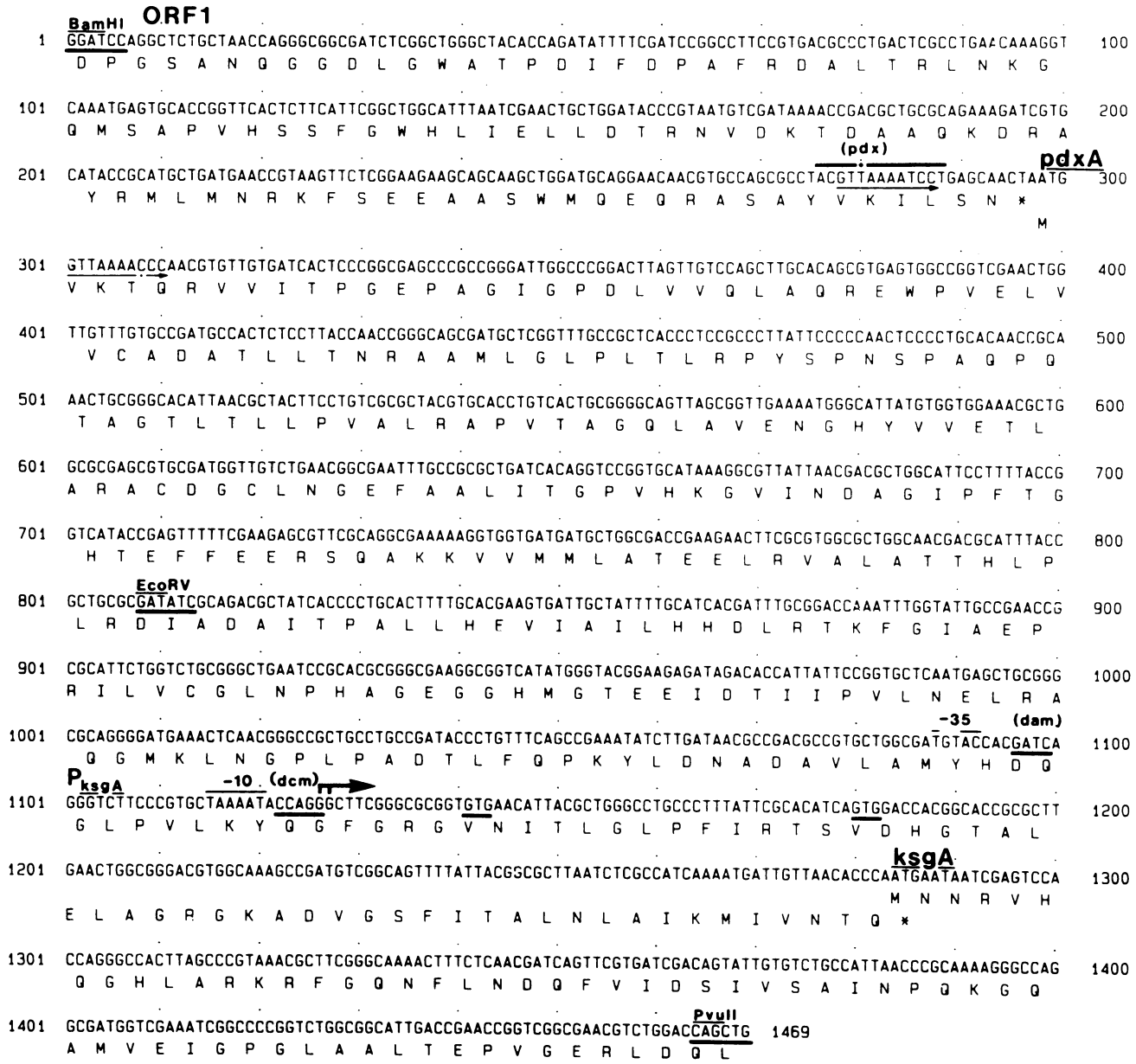

FIG. 3. DNA sequence of *pdxA* and the start of *ksgA*. The sequence extends rightward from the *Bam*HI site shown at the left of Fig. 1. The putative open reading frame (ORF1) upstream of *pdxA* is marked. The sequence and direct repeat that are conserved at the start of the *pdxA* and *pdxB* reading frames are indicated by "pdx" and arrows. P_{ksgA} is shown along with the probable chromosomal transcription initiation point at position 1126 or 1127. Possible translation start points on P_{ksgA} transcripts are underlined (positions 1141 and 1180). Other features are described in the text.

pdxA region. In this coding strand, there were only three long, continuous open reading frames (Fig. 1 and 3). ORF1 extended from the *Bam*HI site to nucleotide (nt) 298 and might encode 11 kDa of the carboxyl terminus of a polypeptide that normally initiates translation upstream of the *Bam*HI site in the bacterial chromosome. The second open reading frame extended from nt 298 to 1287 and encoded a polypeptide with a predicted molecular mass of 35,113 Da (Fig. 3). Substantial evidence (presented below) demonstrated that this open reading frame is *pdxA*. The third open reading frame, which started at nt 1284 and extended beyond the *Pvu*II site at the end of the insert, matched the amino terminus of the *ksgA* gene product (8).

Identification of the second open reading frame as *pdxA* is based on several considerations. First, this was the only reading frame in the coding strand that would be disrupted by insertions into the *Eco*RV site shown in Fig. 1. Second, this open reading frame was the only one that can account for loss of *pdxA* function by the four mini-Mu d1(1734-Km^r) transcriptional fusions and the Km^R(*Eco*RV) insertions described above and in Table 2. Third, minicell translation experiments described below confirmed that (i) the *Bam*HI-*Pvu*II region upstream of *ksgA* (Fig. 1) encodes a polypeptide with a molecular mass of about 36.7 kDa; (ii) expression of this polypeptide is absent from *pdxA*::Km^R(*Eco*RV) constructs; (iii) plasmid pNU167 produces a fusion polypeptide,

as predicted from the cloning junction between the *pdxA* reading frame in Fig. 3 and the pACYC184 vector; and (iv) *pdxA*::mini-Mu d1(1734-Km^r) insertions into pNU167 abolished expression of the fusion polypeptide. Together, these results prove that the open reading frame indicated in Fig. 3 must encode the *pdxA* gene product.

There are two possible translation starts for PdxA, at positions 298 (AUG) and 319 (GUG) (Fig. 3). Presently, it is not possible to tell which of these start codons is actually used, since translation initiation at either one would result in a polypeptide whose size is near that of PdxA on SDS-PAGE gels (see below). Neither candidate was preceded by a good match to the Shine-Dalgarno consensus sequence (38). It is interesting that both possible translation starts are near a sequence and direct repeat, labeled in Fig. 3, which are found at the start of the *pdxB* pyridoxine biosynthesis gene (P. Schoenlein, B. Roa, and M. Winkler, submitted for publication) (see Discussion). If we assume that AUG is the start codon, the PdxA polypeptide contains 329 amino acids. Codon usage in *pdxA* is similar to that in *ksgA*, *apaG*, and *apaH*, and *pdxA* contains a greater number of infrequently and rarely used codons than genes for average *E. coli* "nonregulatory" proteins (data not shown) (2). The predicted *pdxA* polypeptide is probably cytoplasmic and lacks detectable features such as ATP-, NAD-, or zinc-binding sites.

One other feature of the *pdxA* DNA sequence is noteworthy. Previously, the DNA sequence downstream from the *Bgl*I site at position 1075 (Fig. 3) was determined independently by two different groups interested in characterizing *ksgA* (8, 44). There is a discrepancy between these two DNA sequences upstream of the known translation start of *ksgA* (position 1284, Fig. 3). Our DNA sequence in this region agrees with the one reported by Blanchin-Roland et al. (8) and indicates that the stop codon of *pdxA* overlaps the start codon of *ksgA* by 2 bases (Fig. 3). Thus, *pdxA* and *ksgA* must share a common transcript and lie in the same complex operon. Other implications of this gene arrangement are considered in the Discussion.

Expression of PdxA in minicells. In earlier experiments, Blanchin-Roland et al. did not detect a polypeptide that might correspond to the *pdxA* gene product when they expressed a plasmid containing the *Bam*HI-*Eco*RI insert shown in Fig. 1 in maxicells (8). To verify predictions from the DNA sequence analysis, we performed a series of minicell translation experiments with various plasmid constructs. Figure 4A shows polypeptides labeled with [^{35}S]methionine that are encoded by plasmid pUC18 (vector), pNU121 (*pdxA*$^+$ *ksgA*$^+$ *apaG*$^+$ *apaH*$^+$), pNU130 [*pdxA*::Km^R(*Eco*RV)>*ksgA*$^+$ *apaG*$^+$ *apaH*$^+$], or pNU131 [*pdxA*::<Km^R(*Eco*RV) *ksgA*$^+$ *apaG*$^+$ *apaH*$^+$] in minicell strain P678-54. Plasmid pNU121 (*pdxA*$^+$) encoded a 36.7-kDa polypeptide (Fig. 4, lane 2) that was not synthesized from the vector (lane 1) or *pdxA*::Km^R mutant plasmids (lanes 3 and 4). This is exactly the result predicted from the DNA sequence if the 36.7-kDa polypeptide is the *pdxA* gene product. The gradient gel system used in Fig. 4A failed to resolve the *ksgA*, *apaH*, and *bla* gene products; however, a band was detected at the appropriate position for the ApaG polypeptide (lanes 2 to 4).

The putative *pdxA* gene product was expressed at an extremely low level compared with other polypeptides encoded by plasmid pNU121 (Fig. 4A, lane 2). From the DNA sequence, such low-level expression was unanticipated from the methionine content. In plasmid pNU121, the presumed *pdxA* coding region was opposite the P_{lac} promoter; therefore, relatively low PdxA expression possibly reflected lack of efficient transcription. In fact, the band at the bottom of lane 2 in Fig. 4A, which is absent from lanes 3 and 4, was unique to this plasmid construct and probably resulted from P_{lac}-initiated transcription of the *pdxA* region on the DNA strand that normally is not used in the bacterial chromosome.

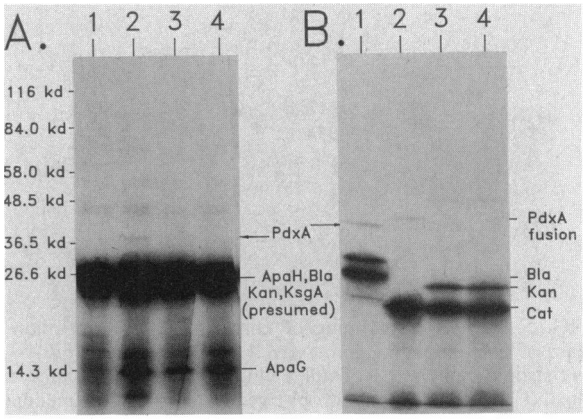

FIG. 4. Expression of plasmid-encoded polypeptides in minicells. Minicells were labeled with [^{35}S]methionine, and polypeptides were resolved on polyacrylamide gradient gels containing SDS as described in Materials and Methods. (A) A 5 to 20% polyacrylamide gradient–SDS gel. Lane 1, pUC18 (control); lane 2, pNU121 (*pdxA*$^+$ *ksgA*$^+$ *apaG*$^+$ *apaH*$^+$ clone in pUC18); lane 3, pNU130 [*pdxA*::Km^R(*Eco*RV)> *ksgA*$^+$ *apaG*$^+$ *apaH*$^+$ derivative of pNU121]; lane 4, pNU131 [*pdxA*::<Km^R(*Eco*RV) *ksgA*$^+$ *apaG*$^+$ *apaH*$^+$ derivative of pNU121]. Positions of molecular mass standards and expected positions of the *apaH*, *bla*, *kan*, *ksgA*, and *apaG* gene products are indicated. The band corresponding to the *pdxA* gene product is marked (see Results). (B) A 12.5% linear polyacrylamide–SDS gel. Lane 1, pNU243 (*pdxA*$^+$ minimal clone derived from pBR322); the PdxA band was absent in pBR322 controls (data not shown) (4); lane 2, pNU167 (*pdxA*$^+$ translational fusion derived from pACYC184); lane 3, pNU209 [*pdxA*::mini-Mu d1(1734-Km^r)-1 derivative of pNU167]; lane 4, pNU210 [*pdxA*::mini-Mu d1(1734-Km^r)-2 derivative of pNU167]. Locations of mini-Mu d1 insertions in plasmids are diagrammed in Fig. 1. The *pdxA*, *bla*, *kan*, and *cat* gene products and the PdxA fusion polypeptide from pNU167 are indicated.

Further minicell translation experiments only partly supported the interpretation that inefficient transcription caused low PdxA expression from plasmid pNU121. In these experiments, the *pdxA* coding region was cloned downstream from known active promoters. Minimal clone pNU243 contains only the *Bam*HI-*Pvu*II fragment, indicated by the thick black line in Fig. 1. In this construct, *pdxA* is downstream from the pBR322 *tet* promoter. The putative 36.7-kDa PdxA polypeptide was again synthesized (Fig. 4B, lane 1), but this time it was expressed at a somewhat greater level compared with other pNU243-encoded polypeptides (compare with Fig. 4A, lane 2). Nevertheless, PdxA expression was still relatively low. Similar results were obtained for plasmid pNU122, in which *pdxA* was cloned downstream from P_{lac} in pUC19 and isopropylthiogalactopyranoside was added to minicell translation reaction mixtures (data not shown). Thus, expression of PdxA is relatively weak in minicells, even when active promoters are upstream from *pdxA*. Although several explanations are possible, this low expression may partially reflect codon usage in *pdxA*, since other polypeptides encoded by genes containing infrequent and rare codons, such as *pdxB* and *hisT*, are also expressed at low levels in minicells (4, 28). Finally, transcription experiments presented in the next section show that the 298-bp

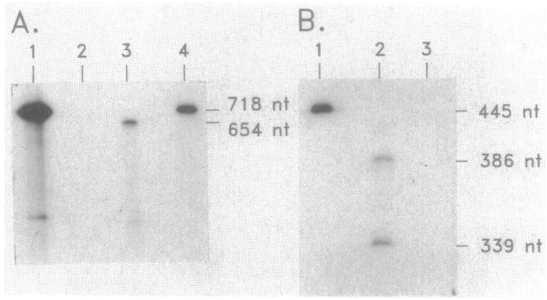

FIG. 5. RNase T_2 mapping of chromosomal transcription upstream of and extending into *pdxA* (A) and from the *pdxA-ksgA* intercistronic region (B). Total cellular RNA was isolated from strain NU426 (W3110 prototroph) grown in LB+Cys medium at 37°C, and 50 μg of total RNA was hybridized to ^{32}P-labeled RNA probes as described in Materials and Methods. (A) *Bam*HI-*Ava*II RNA probes (probe 1, Fig. 1). Lane 1, Control showing RNA with the same sequence as the DNA coding strand (coding-strand RNA probe 1); lane 2, coding-strand RNA probe 1 hybridized with NU426 RNA; lane 3, noncoding-strand RNA probe 1 hybridized with NU426 RNA; lane 4, control showing RNA with the same sequence as the DNA noncoding strand (noncoding-strand RNA probe 1). The positions of noncoding-strand RNA probe (718 nt) and full-length-protected noncoding-strand RNA probe (654 nt) are marked (see text). (B) *Bgl*I-*Pvu*II RNA probes (probe 3, Fig. 1). Lane 1, Noncoding-strand RNA probe 3 control; lane 2, noncoding-strand RNA probe 3 hybridized with NU426 RNA; lane 3, coding-strand RNA probe 3 hybridized with NU426 RNA. The positions of noncoding-strand RNA probe (445 nt), full-length-protected noncoding-strand RNA probe (386 nt), and protected noncoding-strand RNA probe corresponding to transcription initiations from P_{ksgA} (339 nt) are marked (see text).

region upstream of *pdxA* does lack a promoter, which suggests that PdxA expression from pNU121 (Fig. 4A, lane 2) probably resulted from inefficient transcription from adjoining vector DNA.

One last series of minicell translation experiments confirmed that the *pdxA* reading frame indicated in Fig. 3 does indeed encode the 36.7-kDa polypeptide detected in minicells. Based on its DNA sequence, plasmid pNU167, which contains a *Bam*HI-*Hpa*I subfragment 191 bp shorter than the *Bam*HI-*Pvu*II in pNU243 (Fig. 1), should encode a fusion polypeptide in which the last two amino acids of PdxA are replaced by 96 amino acids encoded by an adjoining segment of the pACYC184 vector. This fusion polypeptide should be transcribed from the pACYC184 *tet* promoter. As shown in Fig. 4B, lane 2, plasmid pNU167 did encode a polypeptide of about the expected size (≈45 kDa), which was again expressed at a relatively low level in minicells. Furthermore, mini-Mu dl(1734-Km^r) insertions, which inactivated *pdxA* function on pNU167, prevented expression of the PdxA fusion polypeptide (Fig. 4B, lanes 3 and 4). These results strongly support the contention that *pdxA* is the open reading frame indicated in Fig. 1 that overlaps *ksgA* by 4 bp and encodes a polypeptide with an apparent molecular mass of about 36.7 kDa. In addition, these findings show that the last two amino acids of PdxA can be replaced without loss of function.

Transcription of *pdxA* and *ksgA*. To learn whether *pdxA* is the first gene in the *pdxA-ksgA-apaG-apaH* operon, we examined in vivo chromosomal transcription that extended into *pdxA* from the region upstream (Fig. 5A). We also checked an internal region to verify contiguous transcription of one DNA strand throughout the *pdxA* region (data not shown). Last, we measured transcription from the *pdxA-ksgA* junction to confirm that *pdxA* and *ksgA* are indeed members of the same complex operon and to locate exactly the P_{ksgA} internal promoter, which had been detected by Blanchin-Roland et al. (8). Total cellular RNA was isolated from strain NU426 (*pdxA*$^+$) grown exponentially in LB+Cys medium at 37°C, and RNase T_2 mapping of in vivo transcripts was performed as detailed in Materials and Methods. The RNA probes are diagrammed at the left of Fig. 1, corresponding to (i) the region that starts upstream and extends into *pdxA* (probe 1); (ii) an internal region within *pdxA* (probe 2); and (iii) the *pdxA-ksgA* intercistronic region (probe 3). The DNA strand shown in Fig. 3 is referred to as the coding strand. Sizes of protected RNA probes were determined relative to 10 RNA size standards (data not shown), because we found discrepancies between known sizes of RNA species and predicted sizes based on DNA sequence ladders.

Transcripts originating upstream of and extending into *pdxA* gave full-length protection of noncoding-strand RNA probe 1 (lane 3, Fig. 5A), whereas no protection of coding-strand RNA probe 1 was detected on longer exposures of autoradiographs (lane 2, Fig. 5A) (data not shown). Full-length-protected probe 1 measured 676 nt, compared with its predicted size of 654 nt, and was shorter than the 718-nt starting RNA probe (lanes 1 and 3, Fig. 5A), because the starting probe contained 64 nt of vector-derived "linker" sequence on its end. The slight disparity between the measured and predicted sizes of the full-length-protected probe probably arose because the RNA size standards were not well distributed in that part of the gel (data not shown). Comparable results were obtained with RNA probe 2, which should be internal to *pdxA* (Fig. 1). In this case, full-length protection of noncoding-strand RNA probe 2 was observed, while no protection of coding-strand RNA probe 2 was detected (data not shown).

These results confirm that *pdxA* transcription occurs in vivo on the complement of the DNA strand shown in Fig. 3 from the *Bam*HI site at position 1 to at least the *Bgl*I site at position 1081 in the bacterial chromosome. We also conclude that chromosomal *pdxA* transcription initiates from a promoter (P_{up}, Fig. 1) located at least 298 bp upstream from the start of *pdxA*. The ORF1 open reading frame encoded by this transcript contains likely codons for *E. coli* and ends with a UAA stop codon that overlaps the putative AUG start codon of *pdxA* (Fig. 3). Together, these observations suggest that *pdxA* is not the first gene of this multifunctional operon.

Transcripts from the *pdxA-ksgA* intercistronic region gave two strongly protected RNA bands from noncoding-strand RNA probe 3 (lane 2, Fig. 5B). Again, no protection of the coding-strand control was detected on overexposed autoradiographs (lane 3, Fig. 5B) (data not shown). The measured size of the upper protected band was 389 nt, which matches the predicted size of 386 nt for full-length-protected probe. The lower band measured 339 ± 2 nt and corresponds to transcription initiations from the P_{ksgA} internal promoter. Based on this determination, P_{ksgA} is located at the position shown in Fig. 3, where the 5′ end of the P_{ksgA} transcript is probably a G residue that corresponds to a protected probe length of 340 or 341 nt. To confirm that the lower protected band represents transcription initiations rather than RNA processing, we mapped transcripts from this region in strain NU812, which contains a polar *pdxA*::Km^R(*Eco*RV) insertion upstream of P_{ksgA} (Fig. 1). The presence of the cassette insertion significantly reduced the relative amount of the upper protected species compared with the lower species, as expected for independent transcription initiation from P_{ksgA}

(data not shown). Finally, densitometer tracings of autoradiographs were quantitated and corrected for radioactive content of the protected RNA probes of different lengths. This analysis indicates that steady-state transcription levels from P_{up} and P_{ksgA} are approximately equal in bacteria growing exponentially in LB+Cys medium at 37°C.

DISCUSSION

In this paper, we show that *pdxA*, which encodes a 35.1-kDa polypeptide required for pyridoxine (vitamin B_6) biosynthesis, forms a complex operon with *ksgA*, which encodes the important 16S rRNA modification enzyme m_2^6A methyltransferase. This finding has several implications for pyridoxal phosphate biosynthesis. Because *ksgA* has already been shown to form a complex operon with *apaG* and *apaH*, which mediate the hydrolysis of the alarmone AppppA (8), our results genetically link pyridoxine biosynthesis to rRNA modification and cellular alarmone levels. Previously, we demonstrated that *pdxB*, which encodes a pyridoxine-biosynthetic enzyme, is the first gene in a five-member operon that includes the *hisT* tRNA modification gene (4). At that time, we speculated that other *pdx* biosynthetic genes might be organized into complex operons to allow the cell genetically to interweave coenzyme biosynthesis with other metabolic processes (3). The finding that *serC(pdxF)* (21) and now *pdxA* are members of complex operons is consistent with this hypothesis. In the case of the *pdxA-ksgA-apaG-apaH* operon, pyridoxine biosynthesis, which is central to amino acid metabolism (6), is potentially linked at the gene control level to ribosome function via *ksgA* (43) and AppppA concentration via *apaG* and *apaH*. The function of AppppA is presently not well understood, yet translation and AppppA are related, since AppppA is synthesized as a side reaction of aminoacyl-tRNA synthetases (9, 34), presumably in response to cellular oxidative stress (9).

The *pdxA-ksgA-apaG-apaH* operon is typical of other multifunctional operons in that its organization and expression are complicated (Fig. 1) (7). One aspect of this organization is particularly relevant to speculations about why these four genes are arranged into a transcription unit. *pdxA-ksgA-apaG-apaH* contains at least two internal promoters, P_{ksgA} (Fig. 1, 3, and 5) (8) and P_{apaG} (Fig. 1) (8). Expression of *ksgA* from P_{ksgA} is sufficient to allow full 16S rRNA modification under some growth conditions, as judged by antibiotic resistance (Fig. 2). Therefore, independent expression of *ksgA*, *apaG*, and *apaH* from P_{ksgA} and of *apaG* and *apaH* from P_{apaG} is possible (Fig. 1). However, we also show that a substantial portion of *ksgA* expression (≈50%, Fig. 5B) seems to originate in vivo from the P_{up} promoter shared by *pdxA* and *ksgA* (Fig. 1). Thus, this arrangement of nested promoters could allow coregulation of all operon genes or independent regulation of certain genes depending on cellular metabolic state. In this regard, P_{ksgA} contains *dam* and *dcm* DNA methylation sites and is reminiscent of a stringently controlled promoter (Fig. 3) (14, 37). These conjectures need to be tested in future experiments.

Several other structural features of the *pdxA-ksgA-apaG-apaH* operon are interesting and may be relevant to gene regulation. The putative start codon of *pdxA* is near a nucleotide sequence and direct repeat that are also found close to the translational start of *pdxB* (Fig. 3) (P. Schoenlein, B. Roa, and M. Winkler, submitted for publication). This remarkable conservation suggests that *pdxA* and *pdxB* expression may be regulated and coordinated at the translational level. The potential regulatory sequence and direct repeat are not found in *serC(pdxF)*, but this gene may be regulated differently than *pdxA* and *pdxB*, because it is involved in serine as well as pyridoxine biosynthesis (20; H. Lam and M. Winkler, unpublished observations).

Another unusual feature of the *pdxA-ksgA-apaG-apaH* operon is the overlap or proximity of all the polypeptide reading frames (Fig. 1 and 3). The *pdxA* and *ksgA* reading frames overlap by 4 nt (Fig. 3), while *ksgA* and *apaG* or *apaG* and *apaH* are separated by 2 and 6 nt respectively (8). Results presented above also suggest that *pdxA* may not be the first gene in the operon, which means that ORF1 might encode a gene whose stop codon overlaps the start codon of *pdxA* (Fig. 3). Both *pdxA* and *ksgA* lack identifiable Shine-Dalgarno sequences (Fig. 4) (38), and part of the reason PdxA expression was so low in minicells (Fig. 4) may be lack of coupling to ORF1 translation in plasmid constructs. Moreover, preliminary data suggest that translation of an upstream region is necessary for translation of *ksgA* (44). In this instance, *pdxA* and *ksgA* might be translationally coupled on transcripts from P_{up}, whereas peptides at the end of the *pdxA* reading frame (position 1141 or 1180, Fig. 3) may need to be translated to get *ksgA* expression on transcripts from P_{ksgA}. Finally, Blanchin-Roland et al. suggested that translational coupling also takes place between *apaG* and *apaH* (8). Therefore, by analogy to ribosomal protein operons (32), it is possible that all the genes of the *pdxA-ksgA-apaG-apaH* operon depend on translational coupling for their expression and that translational control of *pdxA* could influence the amounts of the *ksgA*, *apaG*, and *apaH* gene products.

Lastly, in the course of identifying *pdxA*, new chromosomal insertion mutations were constructed that abolished *pdxA* expression (Fig. 1 and 4; Table 2). Complementation of *pdxA*::Km^R mutations, which were constructed in this study, and several *pdxA* point mutations, which had been characterized previously by Dempsey and co-workers (16) by "wild-type" plasmids pNU121, pNU122, and pNU123, but not by their *pdxA*::Km^R(*Eco*RV) mutant derivatives, demonstrated that the "*pdxA* locus" consists of a single complementation group (see Results) (data not shown). Mutants containing *pdxA*::Km^R insertions were not supplemented by compounds that allow partial growth of *pdxB*::Km^R mutants at 30°C (Table 2). Growth of *pdxB*::Km^R mutants on glycoaldehyde, 4-pyridoxic acid, and D-alanine was attributed to an alternative biosynthetic pathway which functions at low temperature in the absence of the *pdxB* gene product (3). Lack of growth of *pdxA*::Km^R "knockout" mutants on compounds other than B_6 vitamers is consistent with the notion that the *pdxA* and *pdxB* gene products catalyze steps in different branches of the pyridoxine-biosynthetic pathway. Because *pdxB* most likely mediates addition of C-5, C-5′, and C-6 and N-1 to the pyridine ring of pyridoxine (3; H. Lam and M. Winkler, unpublished observations), *pdxA* probably mediates addition of C-2 and C-2′ or C-3, C-4, and C-4′ from a pyruvate or glyceraldehyde-3-phosphate precursor, respectively (18, 25). The role of the *pdxB* and *serC(pdxF)* enzymes will be elaborated in another paper, in which we show that the branch of the pyridoxine-biosynthetic pathway mediated by *pdxB* and *serC(pdxF)* is evolutionarily related to the major serine-biosynthetic pathway.

ACKNOWLEDGMENTS

We thank B. Bachmann, M. Casadaban, J. Clark, W. Dempsey, P. Matsumura, B. Nichols, J. Roth, and C. Yanofsky for bacterial

strains and W. Dempsey, H.-M. Lam, B. Nichols, P. Schoenlein, and E. Tancula for discussions and comments.

This work was supported by Public Health Service grant GM37561 from the National Institute of General Medical Sciences.

LITERATURE CITED

1. **Andresson, O. S., and J. E. Davies.** 1980. Genetic organization and restriction enzyme cleavage map of the *ksgA-pdxA* region of the *Escherichia coli* chromosome. Mol. Gen. Genet. **179:**211–216.
2. **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.
3. **Arps, P. J., and M. E. Winkler.** 1987. An unusual genetic link between vitamin B_6 biosynthesis and tRNA pseudouridine modification in *Escherichia coli* K-12. J. Bacteriol. **169:**1071–1079.
4. **Arps, P. J., and M. E. Winkler.** 1987. Structural analysis of the *Escherichia coli* K-12 *hisT* operon by using a kanamycin resistance cassette. J. Bacteriol. **169:**1061–1070.
5. **Bartlett, D. H., and P. Matsumura.** 1984. Identification of *Escherichia coli* region III flagellar gene products and description of two new flagellar genes. J. Bacteriol. **160:**577–585.
6. **Bender, D. A.** 1985. Amino acid metabolism. John Wiley & Sons, Inc., New York.
7. **Björk, G. R.** 1985. *E. coli* ribosomal protein operons: the case of the misplaced genes. Cell **42:**7–8.
8. **Blanchin-Roland, S., S. Blanquet, J.-M. Schmitter, and G. Fayat.** 1986. The gene for *Escherichia coli* diadenosine tetraphosphatase is located immediately clockwise to *folA* and forms an operon with *ksgA*. Mol. Gen. Genet. **205:**515–522.
9. **Bochner, B. R., P. C. Lee, S. W. Wilson, C. W. Cutler, and B. N. Ames.** 1984. AppppA and related adenylylated nucleotides are synthesized as a consequence of oxidation stress. Cell **37:**225–232.
10. **Bognar, A. L., C. Osborne, and B. Shane.** 1987. Primary structure of the *Escherichia coli folC* gene and its folylpolyglutamate synthetase-dihydrofolate synthetase product and regulation of expression by an upstream gene. J. Biol. Chem. **262:**12337–12343.
11. **Castilho, B. A., P. Olfson, and M. J. Casadaban.** 1985. Plasmid insertion mutagenesis and *lac* gene fusion with mini-Mu bacteriophage transposons. J. Bacteriol. **158:**488–495.
12. **Costa, R. H., E. Lai, and J. E. Darnell.** 1986. Transcriptional control of the mouse prealbumin (transthyretin) gene: both promoter sequences and a distinct enhancer are cell specific. Mol. Cell. Biol. **6:**4697–4708.
13. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **de Boer, H. A., S. F. Gilbert, and M. Nomura.** 1979. DNA sequences of promoter regions for rRNA operons *rrnE* and *rrnA* in *E. coli*. Cell **17:**201–209.
15. **Dempsey, W. B.** 1965. Control of pyridoxine biosynthesis in *Escherichia coli*. J. Bacteriol. **90:**431–437.
16. **Dempsey, W. B.** 1969. Characterization of pyridoxine auxotrophs of *Escherichia coli*: P1 transduction. J. Bacteriol. **97:**1403–1410.
17. **Dempsey, W. B.** 1971. Control of vitamin B_6 biosynthesis in *Escherichia coli*. J. Bacteriol. **108:**415–421.
18. **Dempsey, W. B.** 1987. Synthesis of pyridoxal phosphate, p. 539–543. *In* F. C. Neidhardt (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
19. **Dempsey, W. B., and H. N. Christensen.** 1962. The specific binding of pyridoxal 5′-phosphate to bovine serum albumin. J. Biol. Chem. **237:**1113–1120.
20. **Dempsey, W. B., and H. Itoh.** 1970. Characterization of pyridoxine auxotrophs of *Escherichia coli*: serine and *pdxF* mutants. J. Bacteriol. **104:**658–667.
21. **Duncan, K., and J. R. Coggins.** 1986. The *serC-aroA* operon of *Escherichia coli*: a mixed function operon encoding enzymes from two different amino acid biosynthetic pathways. Biochem. J. **234:**49–57.
22. **Hames, B. D.** 1981. An introduction to polyacrylamide gel electrophoresis, p. 1–92. *In* B. D. Hames and D. Rickwood (ed.), Gel electrophoresis of proteins: a practical approach. IRL Press Ltd., London.
23. **Hedrick, J. L., and E. H. Fischer.** 1965. On the role of pyridoxal 5′-phosphate in phosphorylase. I. Absence of classical vitamin B_6-dependent enzymatic activities in muscle glycogen phosphorylase. Biochemistry **4:**1337–1342.
24. **Henikoff, S., and J. C. Wallace.** 1988. Detection of protein similarities using nucleotide sequence databases. Nucleic Acids Res. **16:**6191–6204.
25. **Hill, R. E., and I. D. Spenser.** 1986. Biosynthesis of vitamin B_6, p. 417–476. *In* D. Dolphin, R. Poulson, and O. Avramovic (ed.), Coenzymes and cofactors, vol. I: vitamin B_6 pyridoxal phosphate. John Wiley & Sons, Inc., New York.
26. **Hughes, K. T., and J. R. Roth.** 1984. Conditionally transposition-defective derivative of Mu d1(Ap *lac*). J. Bacteriol. **159:**130–137.
27. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Marvel, C. C., P. J. Arps, B. C. Rubin, H. O. Kammen, E. E. Penhoet, and M. E. Winkler.** 1985. *hisT* is part of a multigene operon in *Escherichia coli* K-12. J. Bacteriol. **161:**60–71.
29. **Mechulam, Y., M. Fromant, P. Mellot, P. Plateau, S. Blanchin-Roland, G. Fayat, and S. Blanquet.** 1985. Molecular cloning of the *Escherichia coli* gene for diadenosine 5′,5‴-*P*1,*P*4-tetraphosphate pyrophosphohydrolase. J. Bacteriol. **164:**63–69.
30. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9:**309–321.
31. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Nomura, M., R. Gourse, and G. Baughman.** 1984. Regulation of the synthesis of ribosomes and ribosomal components. Annu. Rev. Biochem. **53:**75–117.
33. **Nonet, M. L., C. C. Marvel, and D. R. Tolan.** 1987. The *hisT-purF* region of the *Escherichia coli* K-12 chromosome. J. Biol. Chem. **262:**12209–12217.
34. **Plateau, P., M. Fromant, and S. Blanquet.** 1987. Heat shock and hydrogen peroxide responses of *Escherichia coli* are not changed by dinucleoside tetraphosphate hydrolase overproduction. J. Bacteriol. **169:**3817–3820.
35. **Sanger, F., and A. R. Coulson.** 1978. The use of thin acrylamide gels for DNA sequencing. FEBS Lett. **87:**107–110.
36. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
37. **Sarmientos, P., and M. Cashel.** 1983. Carbon starvation and growth rate-dependent regulation of the *Escherichia coli* ribosomal RNA promoters: differential control of dual promoters. Proc. Natl. Acad. Sci. USA **80:**7010–7013.
38. **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.
39. **Snell, E. E., and B. E. Haskell.** 1971. Metabolism of water-soluble vitamins, section C: the metabolism of vitamin B_6, p. 47–71. *In* M. Florkin and E. H. Stotz (ed.), Comprehensive biochemistry. Elsevier, Amsterdam.
40. **Tadera, K., E. Mori, F. Yagi, A. Kobayashi, K. Imada, and M. Imabeppu.** 1985. Isolation and structure of a minor metabolite of pyridoxine in seedlings of *Pisum sativum*. L. J. Nutr. Sci. Vitaminol. **31:**403–408.
41. **Tani, Y., and W. B. Dempsey.** 1973. Glycoaldehyde is a precursor of pyridoxal phosphate in *Escherichia coli* B. J. Bacteriol. **116:**341–345.
42. **Uyeda, K.** 1969. Reaction of phosphofructokinase with maleic anhydride, succinic anhydride, and pyridoxal 5′-phosphate. Biochemistry **8:**2366–2373.
43. **van Buul, C. P. J. J., and P. H. Van Knippenberg.** 1985.

Nucleotide sequence of the *ksgA* gene of *Escherichia coli*: comparison of methyltransferases affecting dimethylation of adenosine in ribosomal RNA. Gene **38**:65–72.

44. **Van Gemen, B., H. J. Koets, C. A. M. Plooy, J. Bodlaender, and P. H. Van Knippenberg.** 1987. Characterization of the *ksgA* gene of *Escherichia coli* determining kasugamycin sensitivity. Biochimie **69**:841–848.
45. **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.
46. **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**:103–109.