JOURNAL OF BACTERIOLOGY, Mar. 1990, p. 1621–1627
0021-9193/90/031621-07$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 172, No. 3

Sequence and Transcriptional Pattern of the Essential *Escherichia coli secE-nusG* Operon

WILLA LEE DOWNING,[1] SUSAN L. SULLIVAN,[2] MAX E. GOTTESMAN,[2] AND PATRICK P. DENNIS[1]*

Department of Biochemistry, 2146 Health Sciences Mall, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5,[1] and Institute of Cancer Research, Columbia University College of Physicians and Surgeons, New York, New York 10032[2]

Received 20 September 1989/Accepted 4 December 1989

Two genes, *secE* and *nusG*, situated between the *tufB* and ribosomal protein *rplKAJL* operons in the *rif* region at 90 min on the *Escherichia coli* chromosome, have been sequenced and characterized. The *secE* gene encodes a 127-amino-acid-long polypeptide, which is an integral membrane protein essential for protein export (P. J. Schatz, P. D. Riggs, A. Jacq, M. J. Fath, and J. Beckwith, Genes Dev. 3:1035–1044, 1989). The *nusG* gene encodes a 181-amino-acid-long polypeptide and is involved in transcription antitermination. The protein product of *nusG* is essential for bacterial viability. The *secE-nusG* genes are cotranscribed, with transcripts initiated at the P_{EG} promoter and terminated at the Rho-independent terminator in the region of the *rplK* promoter. The majority of transcripts are processed at a number of sites in the 5′ untranslated leader region by RNase III and are possibly also processed by a second unidentified nuclease. The role of transcript processing in the regulation of *secE* and *nusG* has not yet been established. The juxtaposition and coregulation of a protein export factor and a transcriptional factor raise questions concerning a functional connection between the two processes.

A cluster of essential genes that encode components of the transcription and translation apparatuses is located at 90 min on the *Escherichia coli* chromosome. The *tufB* operon within the cluster encodes four tRNAs and the translation elongation factor EF-Tu; this operon has been sequenced and its transcription pattern has been partially characterized (1, 28). The *rplKAJL-rpoBC* gene cluster encodes, in the following order, the four 50S subunit ribosomal proteins L11, L1, L10, and L12, and the β and β′ subunits of RNA polymerase; the nucleotide sequence of this region has also been determined, and regulation and transcription of these genes have been studied extensively (5, 10, 14, 21).

Two additional genes, *secE* and *nusG*, have been located between the *tufB* and *rplKAJL* operons (24; Fig. 1). The essential *secE* gene encodes a component of the protein export apparatus and has been shown to be an integral membrane protein (24). The *nusG*-encoded protein, originally detected by in vitro transcription and translation of λ*rif*^{d} (12) or after infection by this phage (29), is involved in transcription antitermination. This protein, along with NusA, NusB, and NusE, is required for λ N-mediated antitermination in an in vitro transcription system (J. Greenblatt, personal communication). In addition, mutations in *nusG* suppress the *E. coli nusA1* and *nusE71* mutations and restore the activity of the N protein of λ (S. L. Sullivan, F. Ward, and M. E. Gottesman, manuscript in preparation). Because of the importance of *secE* and *nusG* gene products in essential cellular functions, we have studied the organization, regulation, transcription, and expression of these genes. In this work, we present the sequence and transcription pattern of the *secE-nusG* operon. We demonstrate that *nusG* is essential for bacterial viability, and we show that *secE*, required for protein export, and *nusG*, implicated in transcription antitermination, are arranged in a bicistronic operon.

* Corresponding author.

MATERIALS AND METHODS

Media and growth conditions. Bacteria were grown exponentially at 32 or 37°C in Luria broth (LB) or M9 minimal salts medium (17) supplemented with glucose (0.2% [wt/vol]), required amino acids (50 μg/ml), thiamine (0.5 μg/ml), and NAD when required (1 μg/ml). Growth was monitored by measuring the A_{460}. Kanamycin and ampicillin were each added to 50 μg/ml when necessary.

Plasmid, phage, and bacterial construction. Standard bacteriological techniques, e.g., transformation and P1 transduction, were as previously described (25). The structures and constructions of plasmids are illustrated and described in Fig. 1.

Bacteriophage λ *int2 c*I857 *nusG*::KAN (YU129) was constructed by a three-step procedure. First, the 2.1-kilobase (kb) *Eco*RI fragment from pSS105 was cloned into the unique *Eco*RI site located in the *b* region of cloning vector λNF1955 (9) yielding YU109 (λ *c*I857 *nin5 S*am100 *nusG*$^{+}$). The kanamycin cassette interrupting the *nusG* gene was introduced into an *nin*$^{+}$ *S*$^{+}$ derivative of YU109 by in vivo genetic recombination with pSS110, selecting for λ kanamycin-resistant transducing phage. Finally, the *int2* mutation was introduced by standard phage crosses.

The bacterial strains used in this study are summarized in Table 1. Construction of bacterial strains SS164 and SS165 was performed as follows: MC4100 was spotted with λ *int2 c*I857 *nusG*::KAN (YU129) at 32°C on tryptone broth plates. A kanamycin-resistant, temperature-sensitive lysogen (SS159) was isolated from the center of the spot, and a λ-resistant variant (SS161) was obtained after challenge with λ *i*^{21}*c*I. Strain SS161 was transformed with plasmid pBR322 or pSS105, selecting for ampicillin resistance to yield strains SS164 and SS165, respectively.

DNA sequencing. Standard recombinant DNA procedures were by the method of Maniatis et al. (15). DNA sequencing was done by the dideoxynucleotide-chain termination method (16, 22, 23). Appropriate DNA fragments were cloned into M13mp18, M13mp19, pUC18, or pUC19 as

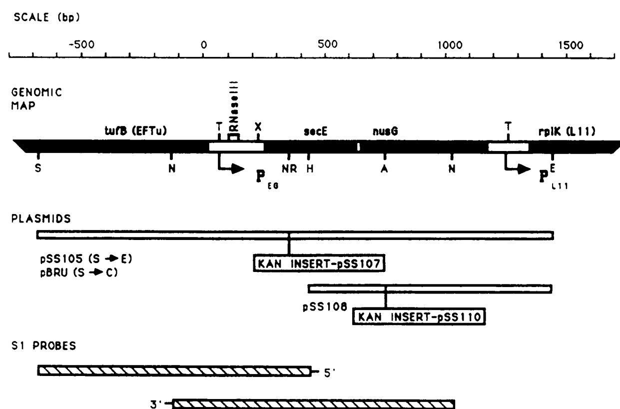
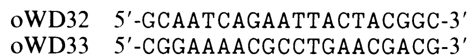


FIG. 1. Genetic organization of the *secE-nusG* gene cluster. The positions of the *tufB* (EF-Tu), *secE*, *nusG*, *rplK* (L11) genes are denoted by the filled rectangles. Abbreviations for selected restriction sites are indicated, and their positions on the nucleotide scale are as follows: *Sma*I (S, −684); *Nci*I (N, −117, 1013); *Nru*I (NR, 347); *Hin*cII, *Hpa*I (H, 419); *Asp*718 (A, 753); *Eco*RI (E, 1438). The *Sma*I site at position −684 corresponds to the *Sma*I site at position 491 by the sequence numbering system of An and Friesen (1). Nucleotide 1158 and the *Eco*RI site at nucleotide 1438 correspond to positions 1 and 280, respectively, in the sequencing numbering system of Post et al. (21). The transcription start sites P_{EG} and P_{L11} (5) are at positions 60 and 1235, respectively. The terminators (T) for *tufB* and *secE-nusG* genes are located at positions 66 and 67 and positions 1238 through 1247, respectively. The RNase III processing sites (RNase III) are situated at nucleotides 96 and 129. A prominent 5′ transcript end which is located at nucleotide 216 is indicated by X. Cloned derivatives of this chromosomal region are as follows. Plasmid pTUB2 contains a 6.4-kb *Eco*RI insert derived from the 90 min region of the *E. coli* chromosome and contains the *rrnB* operon, as well as *tufB*, *secE*, *nusG*, and the proximal portion of *rplK* (18; this work). The *Sma*I-*Eco*RI 2.1-kb fragment from pTUB2 was ligated to an *Eco*RI linker at the *Sma*I end and was inserted into the *Eco*RI site of pBR322 to produce pSS105. Plasmid pBRU is identical to pSS105 except that the *Sma*I-*Eco*RI fragment was obtained from λ *rif*[d]18 (13); the fragment was inserted between the *Eco*RI site and the blunt-ended *Cla*I site (filled in with Klenow enzyme) of pBR322. Plasmid pSS107 was constructed from pSS105 by insertion of a kanamycin resistance cassette (KAN) from pUC4KISS (Pharmacia) into the *Nru*I site at nucleotide 347. The *Hin*cII-*Eco*RI 1.0-kb fragment was inserted into the *Sma*I-*Eco*RI site of pT7-6, a derivative of the T7 promoter vector pT7-1 (26), to give plasmid pSS108. Plasmid pSS110 was obtained by inserting KAN into pSS108 at the *Asp*718 site at position 753. The probes used for S1 nuclease protection experiments were the 5′-end-labeled *Sma*I-*Hpa*I 1.1-kb fragment (top) and the 3′-end-labeled *Nci*I-*Nci*I 1.1-kb fragment (bottom). bp, Base pairs.

templates. Universal forward and reverse primers were used to sequence both strands. The molecular length sequencing ladders used to size the products of primer extension reactions were generated with the oligonucleotide primers oWD32 or oWD33.

Oligonucleotides. Oligodeoxyribonucleotides were synthesized on an Applied Biosystem 380A DNA synthesizer and were deprotected and purified as described by Atkinson and Smith (3). The oligonucleotides used for primer extension were:



oWD32 5′-GCAATCAGAATTACTACGGC-3′
oWD33 5′-CGGAAAACGCCTGAACGACG-3′

oWD32 is complementary to a sequence in the *secE* gene (positions 378 to 397); oWD33 is complementary to a sequence in the proximal region of the *nusG* gene (positions 654 to 673).

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Description[a]
Strain	
C600	*thr leu trp thi recA*
PD828	C600(pBRU)
PD858	C600(pSS107)(=pBRU::KAN)
N99	F⁻ *galK2* Str^r
N2076	F⁻ *thi argH1 nad-84 lacY1 gal-6 nalA1* λ^r *xyl-7 ara-13 mtl-2 str-9 tonA2 rnc*$^+$
N2077	F⁻ *thi argH1 nad-84 lacY1 gal-6 nalA1* λ^r *xyl-7 ara-13 mtl-2 str-9 tonA2 rnc-105*
N3431	HfrPO1 *rel-1 thi-1 lacZ43 rne-3071*(Ts)
N3433	HfrPO1 *rel-1 thi-1 lacZ43 rne*$^+$
MC4100	F⁻ *araD139* Δ(*argF-lac*)*U169 rpsL150 relA1 flbB5301 deoC1 ptsF23 rbsR*
SS77	N99(pSS105)
SS159	MC4100 λ *int2 c*I857 *nusG*::KAN
SS161	MC4100 λ *int2 c*I857 *nusG*::KAN λ^r
SS164	MC4100 λ *int2 c*I857 *nusG*::KAN(pBR322)
SS165	MC4100 λ *int2 c*I857 *nusG*::KAN(pSS105)
SS200	MC4100(pBR322)
Plasmid	
pBRU	*secE nusG*
pSS105	*secE nusG*
pSS107	*secE*::KAN *nusG* (=pBRU::KAN)
pSS108	*nusG*
pSS110	*nusG*::KAN

[a] For the bacterial gene content of the plasmids, only the full-length bacterial genes are listed. For additional details concerning plasmid construction and content, see legend to Fig. 1.

Oligonucleotides (250 ng) were 5′ end labeled at 37°C for 40 min with 10 U of T4 polynucleotide kinase and 100 μCi of [γ-^{32}P]ATP in 20 μl of kinase buffer (0.1 M Tris chloride [pH 8.0], 5 mM dithiothreitol, 10 mM $MgCl_2$). The reaction was terminated by the addition of 1 μl of 0.5 M EDTA (pH 8.0) and incubation at 65°C for 5 min. Eight micrograms of carrier tRNA was added, and the reaction volume was taken up to 100 μl with TE (10 mM Tris chloride [pH 7.5], 1 mM EDTA). The labeled oligonucleotide was purified by two successive ethanol precipitations in the presence of 2.5 M ammonium acetate and was redissolved in 20 to 50 μl of TE.

S1 nuclease mapping. Total cellular RNA was prepared by the method of Downing and Dennis (5). Isogenic *rne*$^+$/*rne* strains N3433 and N3431 were grown in M9 minimal salts media at 30°C to an A_{460} of approximately 0.4. The cultures were then shifted to 44°C for 15 or 30 min; RNA was prepared from these cultures.

The 3′ and 5′ ends of in vivo mRNA transcripts were analyzed by S1 nuclease mapping as described by Berk and Sharp (4) and as modified by Favaloro et al. (6). Conditions for hybridization and S1 nuclease digestion were described previously (5). The 5′-end-labeled *Sma*I-*Hpa*I 1.1-kb fragment and the 3′-end-labeled *Nci*I-*Nci*I 1.1-kb fragment were used to locate 5′ and 3′ transcript ends, respectively (Fig. 1). Both probes span the *tufB-secE* intergenic region. Molecular length standards were *Msp*I fragments of pBR322, 3′ end labeled with Klenow enzyme and [α-^{32}P]dCTP.

Primer extension. Analysis of transcript 5′ ends by the primer extension method was carried out by the method of Newman (19). Ten micrograms of total cellular RNA and 1 ng of 5′-end-labeled oligonucleotide primer were heated at 65°C for 5 min in 10 μl of 160 mM KCl, 40 mM Tris chloride (pH 8.5), and 1 mM EDTA (pH 8.0). The mixture was cooled gradually to 42°C and incubated at 42°C for 1 h. Five units each of avian myeloblastosis virus reverse transcriptase

(Pharmacia, Inc.) and RNase inhibitor (Pharmacia) were then added to the reaction, with 10 μl of 10 mM $MgCl_2$, 10 mM β-mercaptoethanol, and 1 mM of each deoxyribonucleotide. Incubation was continued at 42°C for 1 h. The reaction was stopped by the addition of 2 μl of 0.5 M EDTA and 78 μl of TE. The products were precipitated with ethanol in the presence of 0.3 M sodium acetate. The pellet was dissolved in 5 μl of formamide sequencing dye mix, and the radioactivity was measured by Cerenkov counting. Reaction products were analyzed on an 8% polyacrylamide-urea sequencing gel alongside a sequencing ladder generated with an appropriate single-stranded template and the same primer (but unlabeled) as that used in primer extension.

RESULTS AND DISCUSSION

A physical map of the 1,318-nucleotide-long region between the end of the *tufB* gene and the beginning of the *rplK* gene is depicted in Fig. 1. The complete nucleotide sequence of this region was determined with the *SmaI-EcoRI* 2.1-kb fragments obtained from both genomic DNA (carried on pSS105) and λ*rif*^{d}18 (13). The two sequences were identical (Fig. 2). The region contains two long open reading frames that have been designated *secE* and *nusG*.

The distance between the end of the *tufB* gene and the beginning of the *secE* open reading frame is 229 nucleotides (Fig. 2). Overlapping sequences characteristic of a Rho-independent transcription terminator and an RNA polymerase promoter recognition sequence occur immediately after the *tufB* gene (between nucleotide positions 20 and 70). If functional, the terminator would reduce or prevent extension of the abundant *tufB* transcripts into the *secE-nusG* region. Transcripts initiated at the promoter would contain a 5′ untranslated leader of approximately 180 nucleotides in length.

The *secE* gene (positions 240 to 620) encodes a 127-amino-acid-long polypeptide that is rich in hydrophobic residues. On the basis of a number of different alkaline phosphatase fusions to *secE*, Schatz et al. (24) have suggested that the *secE* gene product is an integral membrane protein containing three membrane-spanning domains. These domains, representing residues 19 through 36, 45 through 63, and 93 through 111, are 18 or 19 amino acids in length and are devoid of charged residues. The amino terminus of the protein is believed to be localized to the inside surface and the carboxy terminus to the outside surface of the cell membrane. The position of the initiating methionine codon at nucleotide 240 is supported by the isolation of a *secE-phoA* gene fusion with a junction immediately after the GAA glutamic acid codon at nucleotide 282. The initiation codon is preceded by a ribosome binding sequence at positions 230 to 233. The isolation of a cold-sensitive mutant indicates that the *secE* product is required for cell viability; presumably, it plays an essential role in the bacterial protein translocation system (24).

Only a single nucleotide separates *secE* from the open reading frame designated *nusG*. The *nusG* gene begins with an ATG methionine codon at position 625 and encodes a polypeptide of 181 amino acids in length. This protein contains a high proportion of acidic (14%) and basic (15%) residues and therefore is probably not an integral membrane protein. Two fusions of alkaline phosphatase to *nusG* at codons three and six confirm the position of the initiation codon (24). This conclusion has recently been substantiated by an N-terminal amino acid sequence of the purified nusG protein. (J. Greenblatt, personal communication).

The *nusG-rplK* intergenic space is 158 nucleotides in length. This region contains the major promoter for transcription of the *rplKAJL-rpoBC* gene cluster, which initiates at or near nucleotide 1235 (Fig. 2) (5, 21). This promoter region overlaps the terminator site for transcripts exiting the *nusG* gene.

The *nusG* gene is essential. We have demonstrated that *nusG* is essential for bacterial viability in a series of gene disruption experiments. The bacteriophage λ *c*I857 *int*2 *nusG*::KAN (YU129) carries a *secE-nusG* region disrupted by the insertion of a kanamycin resistance cassette at the *Asp*718 restriction site (nucleotide 753) within the *nusG* coding region. Lysogens of MC4100 were selected as kanamycin-resistant survivors of infection with YU129 at 32°C (see Materials and Methods; Fig. 3). Since YU129 is integration defective, stable lysogens arose almost exclusively through homologous recombination in or around the *secE* and *nusG* genes and were merodiploid for the region *nusG*$^{+}$-λ *int*2 *c*I857-*nusG*::KAN. Lysogen SS159 was first rendered λ resistant and was subsequently transformed with either pBR322 or pSS105 (*nusG*$^{+}$) to give strains SS164 and SS165, respectively.

Lysogens for λ *c*I857 are killed by thermal induction of the prophage; rare temperature-resistant survivors represent cells in which homologous recombination between flanking sequences removed the prophage prior to induction. If the intact chromosomal *nusG* gene is dispensible for growth, a significant percentage of the surviving bacteria should carry the *nusG*::KAN marker. However, if the chromosomal *nusG* gene is essential, virtually all of the survivors should be kanamycin sensitive. Our results support the latter possibility. Strains SS164 and SS165 were plated at 42°C on LB-ampicillin and LB-kanamycin plates, and the fraction of survivors that were kanamycin resistant for each strain was determined. Whereas 12.6% of the SS165 survivors were resistant to kanamycin, only 0.02% of the SS164 survivors grew on LB-kanamycin (Table 2).

The kanamycin-resistant survivors of strain SS165 presumably carry a defective *nusG*::KAN allele on the chromosome and a complementing *nusG* allele on the plasmid. To support this assumption, we attempted to P1 transduce the *nusG*::KAN marker from the temperature-resistant SS165 survivors to a new background. Numerous kanamycin-resistant transductants of strain SS77 with a plasmid-borne *nusG*$^{+}$ allele were obtained. In contrast, no kanamycin-resistant transductants of strain SS200 carrying pBR322 were obtained. These data indicate that a *nusG* null mutation cannot be introduced into *E. coli* in the absence of a complementing *nusG*$^{+}$ allele. Taken together with the previous results of Schatz et al. (24), it is clear that both the *secE* and *nusG* genes are essential for viability.

Transcript mapping. Plasmids pSS105 and pBRU contain the 2.1-kb *SmaI-EcoRI* fragment and are capable of complementing lethal mutations in the chromosomal *secE* (24) and *nusG* genes. Neither plasmid contains the upstream *tufB* promoter, suggesting the *secE* and *nusG* are transcribed independently of *tufB*. In vivo transcripts derived from the *secE-nusG* region on the bacterial chromosome and the plasmid pBRU were characterized by primer extension and S1 nuclease protection analysis. For this purpose, two synthetic oligonucleotides, one complementary to a region in *secE* (oWD32) and the other complementary to a region in *nusG* (oWD33), were prepared.

The 5′ transcript end sites in the *tufB-secE* intergenic space were analyzed with oWD32 to prime reverse transcription with total RNA isolated from a number of different

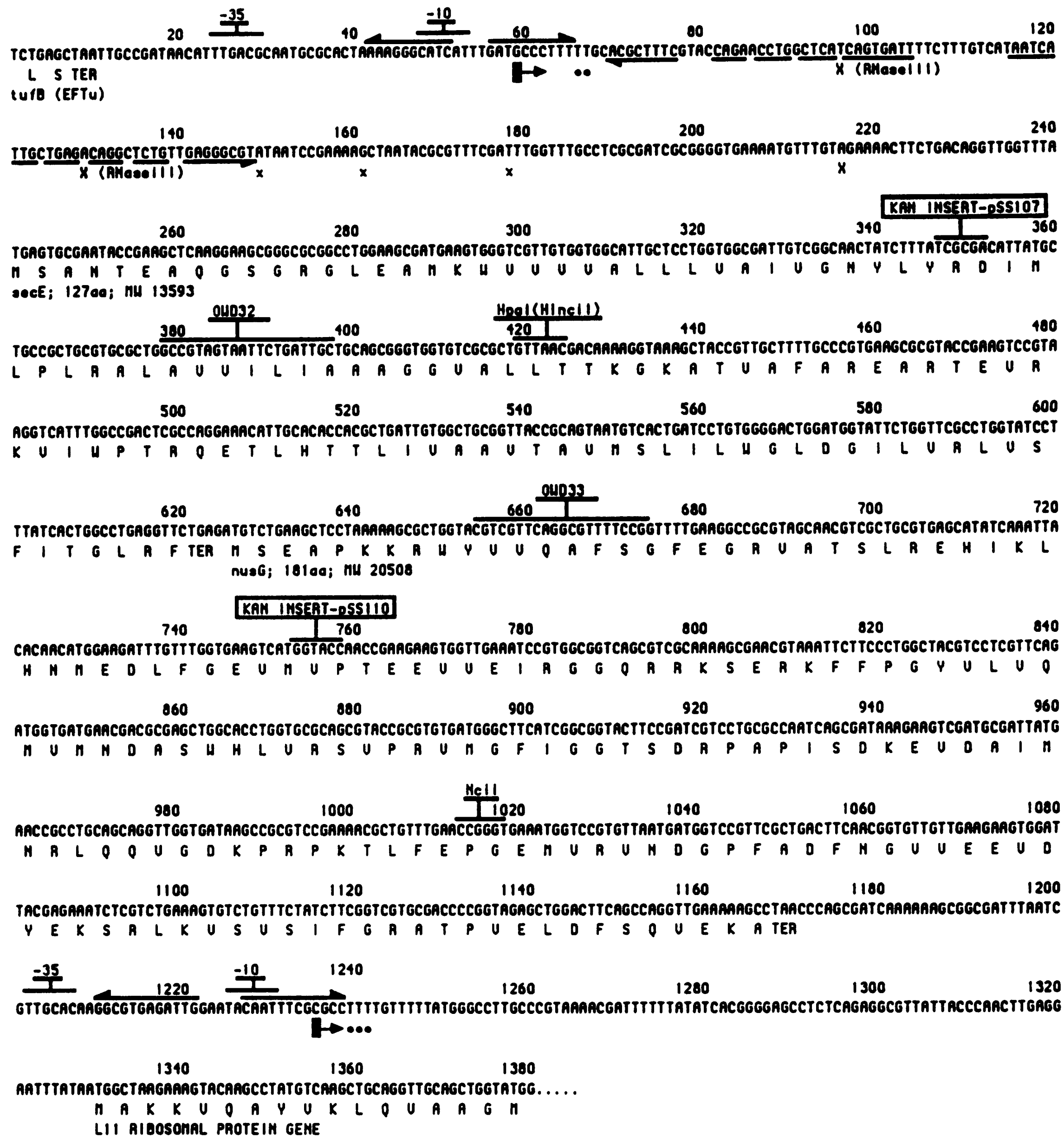


FIG. 2. Nucleotide sequence of *secE-nusG* genes. The predicted amino acid sequences of *secE* and *nusG* are given below the DNA sequence. The *secE* gene is located between nucleotides 240 and 620; the *nusG* gene is located between nucleotides 625 and 1167. The P_{EG} and P_{L11} transcription initiation sites are depicted by arrows at positions 59 and 1235, respectively. The −10 and −35 sequences associated with these 5′ transcript end sites are indicated. Other 5′ transcript ends that originate from processing or weak promoters are indicated by X for major and x for minor mRNA species. The sites of RNase III processing are noted. Sites of transcription termination of *tufB* mRNA and the *secE-nusG* mRNA are shown by filled circles at positions 66 and 67 and 1239 through 1241, respectively. Sequences exhibiting inverted repeat symmetry associated with termination are indicated by overbars. Oligonucleotides oWD32 and oWD33, which were used as primers for primer extension experiments, are complementary to the indicated sequences. The kanamycin resistance cassette was inserted in the *Nru*I site at position 347 (pSS107) or in the *Asp*718 site at position 753 (pSS110). The *Hpa*I site (419) and the *Nci*I site (1013) indicate the ends of restriction fragments, *Sma*I-*Hpa*I (1.1 kb) and *Nci*I-*Nci*I (1.1 kb), respectively, which were used as probes for S1 nuclease mapping. The *Hin*cII site (419) is the same as the *Hpa*I site. The terminal portion of the nucleotide sequence of An and Friesen (1) extends to position 79. The two sequences are identical in the overlapping region with one exception; beginning at postion 41, our nucleotide sequence has a run of four consecutive As as compared with a run of three consecutive As in the sequence of An and Friesen (1). These data have been submitted to GenBank under accession no. M30610.

bacterial strains. A total of seven different 5′-end sites were evident with RNA from strain C600 (Fig. 4A, lanes 3 and 5); these sites are located at or near nucleotide positions 60, 96, 129, 149, 161, 178, and 216. Only the sites at positions 60 and 161 are preceded by easily recognizable and appropriately spaced −10 and −35 promoter consensus sequences. The intensities of the seven 5′-end sites were uniformly enhanced when RNA from strain C600 containing the pBRU plasmid (PD828) was used as the template (Fig. 4A, lane 1). This observation indicates that the transcripts derived from the *secE-nusG* region of the chromosome and the recombinant plasmid are identical and implies that transcription is not

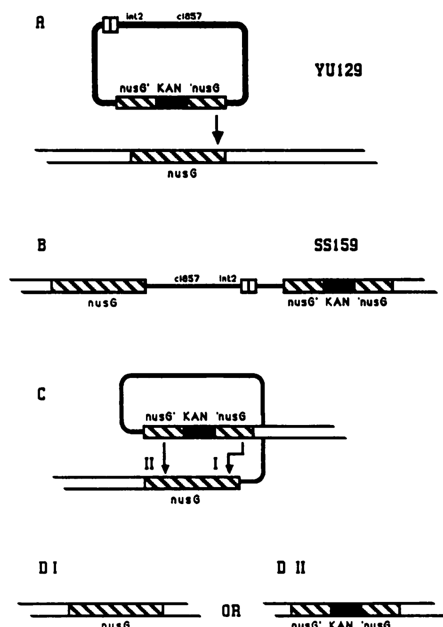


FIG. 3. Disruption of the *nusG* gene. (A) λ *int2 cI857 nusG*::KAN (YU129) inserts into the *E. coli* chromosome by homologous recombination. (B) The resultant *nusG*-λ-*nusG*::KAN merodiploid. (C) Recombination between the *nusG* repeats removes the λ prophage. (DI and DII) Two possible nonlysogenic recombination products.

dependent on the upstream *tufB* promoter. The oligonucleotide oWD33, complementary to a region in *nusG*, was also used to locate 5′ transcript ends. The 5′-end sites of the products generated with this primer correspond to those generated with oWD32 (data not shown).

Plasmid pSS107 contains a kanamycin cassette inserted into the *NruI* site at nucleotide position 347 within the *secE* gene. When RNA from a strain carrying this plasmid (PD858) was used in the primer extension assay with oWD32 as primer, only the low-level transcripts derived from the chromosomal *secE-nusG* region were detected (Fig. 4A, lanes 2 and 4). S1 nuclease analysis clearly demonstrated that few, if any, transcripts exit from the kanamycin cassette (data not shown). Together, these results indicate that the *secE* and *nusG* genes are cotranscribed and that the kanamycin cassette on plasmid pSS107 induces transcriptional polarity on the downstream *nusG* gene.

The two 5′-end sites, at nucleotide positions 96 and 129, are located at nearly opposite positions within a region of inverted repeat symmetry. The endonuclease RNase III is known to recognize interrupted RNA hairpins and to cleave

TABLE 2. A *nusG*::KAN mutant is inviable

Strain	Plasmid	Survivors per ml (10^4)		Kan^r/Amp^r (%)
		LB-amp	LB-kan	
SS164	pBR322	1540	0.3	0.02
SS165	pSS105 (*nusG*$^+$)	460	57.9	12.6

[a] Strains SS164 and SS165 are λ^r lysogens of λ *cI857 nusG*::KAN, bearing pBR322 or pSS105 (*nusG*$^+$), respectively. Cells were grown overnight in LB-ampicillin (amp) (50 μg/ml) at 32°C to a density of 4×10^9 cells per ml. Appropriate dilutions were plated on LB-ampicillin (50 μg/ml) or LB-kanamycin (kan) (50 μg/ml), and the number of colonies were determined after overnight incubation at 42°C.

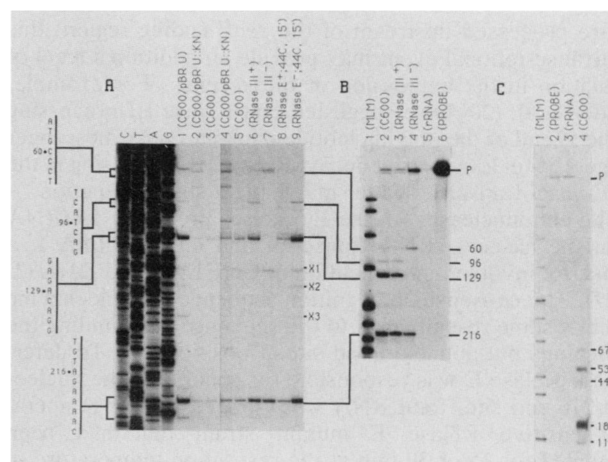


FIG. 4. Transcript mapping by primer extension and S1 nuclease protection. (A) Primer extension with oWD32 as the primer. Reaction products from primer extension experiments were analyzed on an 8% polyacrylamide-urea sequencing gel alongside a sequencing ladder (G, A, T, C). The major 5′ transcript ends are located on the DNA sequence at positions 60, 96, 129, and 216. The minor 5′ mRNA ends are indicated by X1, X2, and X3, and correspond to nucleotide positions 149, 161, and 178, respectively. Ten micrograms of total cellular RNA, prepared from the following strains, was used for each reaction: PD828 [C600(pBRU)] (lane 1), PD858 [C600(pSS107) (= pBRU::KAN)] (lanes 2 and 4), C600 (lanes 3 and 5); N2076 (*rnc*$^+$) (lane 6), N2077 (*rnc*) (lane 7), N3433 (*rne*$^+$) (lane 8), and N3431 (*rne*) (lane 9). Lanes 1 through 3 are short exposures (30 h) and lanes 4 through 9 are long exposures (2 weeks). (B) Nuclease S1 mapping of 5′ transcript ends derived from the *tufB-secE* intergenic region. The 5′-end-labeled *SmaI-HpaI* 1.1-kb restriction fragment was used as the probe; five micrograms of RNA was used in each reaction. Lanes: 1, molecular length markers (MLM) are 3′-end-labeled *MspI* fragments of pBR322 (their lengths are 623, 528, 405, 310, 243, 239, 218, 202, 191, 181, and 161 nucleotides); 2, C600 RNA; 3, N2076 RNA (*rnc*+); 4, N2077 RNA (*rnc*); 5, rRNA (control); 6, 5′-end-labeled *SmaI-HpaI* 1.1-kb probe. The predominant 5′ transcript ends are correlated with their respective primer extension counterparts. The probe (P) and the nucleotide positions of transcript termini are indicated. (C) Nuclease S1 mapping of 3′ transcript ends derived from the *tufB-secE* intergenic region. The 3′-end-labeled *NciI-NciI* 1.1-kb DNA fragment was used as the probe (lane 2). The lane designations are similar to those described for panel B. The 3′ transcript ends are situated on the DNA sequence at positions 11, 18, 44, 53, and 67.

a number of different precursor mRNA and rRNA substrates (5, 7, 11). The RNAs from an RNase III mutant strain (N2077) and its isogenic wild-type parent (N2076) were examined by primer extension to determine if these sites were generated by RNase III cleavage (Fig. 4A, lanes 6 and 7). In the mutant strain, the 5′-end sites at positions 96 and 129 were greatly reduced and the intensity of the 5′-end site at position 60 was correspondingly increased. This result suggests that a precursor RNA with a 5′-end site at position 60 is either partially or slowly cleaved by RNase III at position 96 or 129 or both and that the site at position 60 probably represents the major transcription initiation site for the *secE-nusG* mRNA.

As implied by their relative autoradiogram intensities, transcripts with 5′ termini at positions 129 and 216 are the predominant mRNA species (Fig. 4A and B). The 5′ transcript end at nucleotide 216 is of unknown origin; it may have been generated from the transcript initiated at nucleotide 60 by some unidentified nucleolytic activity. If so, it appears that the majority of transcripts initiated at position

60 are processed upstream of the *secE* coding region; this posttranscriptional event may provide an additional level of regulation in the expression of these genes. For example, Portier et al. (20) have suggested that RNase III processing at the 5′ end of the polynucleotide phosphorylase messenger triggers 5′-to-3′ transcript decay. RNase III processing in the *secE-nusG* transcript leader may have a similar function.

The endonuclease, RNase E, excises precursor 5s rRNA from the nascent rRNA transcript and cleaves RNA I, a transcript involved in replication of ColE1 plasmid DNA (2, 8, 27). The consensus recognition sequence for this enzyme exhibits some resemblance to the sequence surrounding the anomalous but abundant end site at position 216. To determine if RNase E was responsible for generating the nucleotide 216-end site, total RNA was isolated from a temperature-sensitive RNase E mutant strain that had been incubated for 15 or 30 min at the restriction temperature of 44°C. Primer extension with oWD32 indicated that none of the extraneous 5′-end sites, including the one at position 216, were produced by RNase E cleavage (Fig. 4A, lanes 8 and 9).

The 5′ transcript ends detected by primer extension were confirmed by S1 nuclease protection experiments (Fig. 4B). Total cellular RNAs isolated from strain C600 and rnc^+/rnc strains N2076 and N2077 were used to protect the 5′-end-labeled *Sma*I-*Hpa*I 1.1-kb fragment spanning the *tufB-secE* intergenic region (Fig. 1). The ends of fragments protected from S1 nuclease digestion correspond to the transcript ends observed in the primer extension experiments (Fig. 4A and B).

S1 nuclease protection experiments, with the 3′-end-labeled *Nci*I-*Nci*I 1.1-kb probe (Fig. 1), indicated that transcripts exiting the *tufB* gene are efficiently terminated. The longest transcripts terminate at or near nucleotide position 67. This end site is probably a Rho-independent transcription terminator; it lies within a tract of T residues and is preceded by inverted repeat symmetry. A number of other shorter but more abundant transcripts with 3′-end sites near nucleotide positions 11, 18, 44, and 53 were also detected (Fig. 4C). All of these sites lie beyond the *tufB* termination codon. It is unclear whether these 3′-end sites are generated by termination events or by nuclease cleavage in the 3′ untranslated portion of *tufB* mRNA.

The 3′ end of the *secE-nusG* transcript has been previously mapped from nucleotides 1238 to 1247 within a T-tract sequence that is preceded by inverted repeat symmetry (Fig. 2) (5). There is little if any transcription readthrough into the downstream *rplK* gene. The major promoter for the *rplKAJL-rpoBC* gene cluster initiates transcription at or near nucleotide 1235. The overlap of this transcription start site with the *secE-nusG* termination site may permit some regulatory interaction between these two *secE-nusG* and *rplKAJL* gene clusters (5).

In conclusion, *secE* and *nusG* genes have been shown to be essential and their nucleotide sequences have been determined. This completes the nucleotide sequence of the entire *rif* region of the *E. coli* chromosome. The two genes are cotranscribed, with transcription initiation occurring at the P_{EG} promoter and termination occurring at the Rho-independent terminator in the vicinity of the *rplK* (P_{L11}) promoter. The majority of transcripts are processed in the 5′ untranslated leader region by RNase III and probably also by a second unidentified nuclease. Whether transcript processing is a regulatory feature of *nusG* and *secE* expression requires further investigation. In addition, the juxtaposition and coregulation of a transcriptional factor and a protein export factor raise questions concerning a possible functional connection between these two essential cellular processes.

ACKNOWLEDGMENTS

We thank David Apirion for providing strains N3431 and N3433.

This work was supported in part by Public Health Service grant GM37219-03 from the National Institutes of Health and by grant MT6340 from the Medical Research Council of Canada. P.P.D. is a fellow of the Canadian Institute of Advanced Research.

LITERATURE CITED

1. **An, G., and J. D. Friesen.** 1980. The nucleotide sequence of *tufB* and four nearby tRNA structural genes of *Escherichia coli*. Gene **12**:33–39.
2. **Apirion, D., and A. B. Lassar.** 1978. A conditional lethal mutant of *Escherichia coli* which affects the processing of ribosomal RNA. J. Biol. Chem. **253**:1738–1742.
3. **Atkinson, T., and M. Smith.** 1984. Solid phase synthesis of oligodeoxyribonucleotides by the phosphite-triester method, p. 35–81. *In* M. J. Gait (ed.), Oligonucleotide synthesis. A practical approach. IRL Press, Oxford.
4. **Berk, A. J., and P. A. Sharp.** 1978. Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. USA **75**:1274–1278.
5. **Downing, W. L., and P. P. Dennis.** 1987. Transcription products from the *rplKAJL-rpoBC* gene cluster. J. Mol. Biol. **194**:609–620.
6. **Favaloro, J., R. Treisman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
7. **Gegenheimer, P., and D. Apirion.** 1981. Processing of procaryotic ribonucleic acid. Microbiol. Rev. **45**:502–541.
8. **Ghora, B. K., and D. Apirion.** 1979. Identification of a novel RNA molecule in a new RNA processing mutant of *Escherichia coli* which contains 5S rRNA sequences. J. Biol. Chem. **254**: 1951–1956.
9. **Hui, I., K. Maltman, R. Little, S. Hastrup, M. Johnsen, N. Fiil, and P. Dennis.** 1982. Insertions of transposon Tn*5* into ribosomal protein RNA polymerase operons. J. Bacteriol. **152**: 1022–1032.
10. **Jinks-Robertson, S., and M. Nomura.** 1987. Ribosomes and tRNA, p. 1358–1385. *In* F. C. Neidhardt (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
11. **King, T. C., R. Sirdeskmukh, and D. Schlessinger.** 1986. Nucleolytic processing of ribonucleic acid transcripts in procaryotes. Microbiol. Rev. **50**:428–451.
12. **Kirschbaum, J. B., J. Greenblatt, B. Allet, and J. D. Rochaix.** 1976. Studies with a λ*rif* transducing phage carrying the genes for the β and β subunits of E. coli RNA polymerase, p. 503–518. *In* R. Losick and M. Chamberlin (ed.), RNA polymerase. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Kirschbaum, J. B., and E. B. Konrad.** 1973. Isolation of a specialized lambda transducing bacteriophage carrying the beta subunit gene for *Escherichia coli* ribonucleic acid polymerase. J. Bacteriol. **116**:517–526.
14. **Lindahl, L., S. R. Jaskunas, P. P. Dennis, and M. Nomura.** 1975. A cluster of genes in *Escherichia coli* for ribosomal proteins, ribosomal RNA's and RNA polymerase subunits. Proc. Natl. Acad. Sci. USA **72**:2743–2747.
15. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–79.
17. **Miller, J.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Miyajima, A., M. Shibuya, and Y. Kaziro.** 1979. Construction and characterization of the two hybrid ColE1 plasmids carrying *Escherichia coli tufB* gene. FEBS Lett. **102**:207–210.
19. **Newman, A.** 1987. Specific accessory sequences in *Saccharomyces cerevisiae* introns control assembly of pre-mRNAs into spliceosomes. EMBO J. **6**:3833–3839.

20. **Portier, C., L. Dondon, M. Grunberg-Manago, and P. Regnier.** 1987. The first step in the functional inactivation of the *Escherichia coli* polynucleotide phosphorylase messenger is a ribonuclease III processing at the 5′ end. EMBO J. **6:**2165–2170.
21. **Post, L. E., G. D. Strycharz, M. Nomura, H. Lewis, and P. P. Dennis.** 1979. Nucleotide sequence of the ribosomal protein gene cluster adjacent to the gene for RNA polymerase subunit β in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **76:**1697–1701.
22. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143:**161–178.
23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
24. **Schatz, P. J., P. D. Riggs, A. Jacq, M. J. Fath, and J. Beckwith.** 1989. The *secE* gene encodes an integral membrane protein required for protein export in *E. coli*. Genes Dev. **3:**1035–1044.
25. **Silhavy, T., M. Berman, and L. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82:**1074–1078.
27. **Tomcsanyi, T., and D. Apirion.** 1985. Processing enzyme ribonuclease E specifically cleaves RNAI, an inhibitor of primer formation in plasmid DNA synthesis. J. Mol. Biol. **185:**713–720.
28. **Van Delft, J. H. M., D. S. Schmidt, and L. Bosch.** 1987. The tRNA-*tufB* operon transcription termination and processing upstream from *tufB*. J. Mol. Biol. **197:**647–657.
29. **Yamamoto, M., and M. Nomura.** 1979. Organization of genes for transcription and translation in the *rif* region of the *Escherichia coli* chromosome. J. Bacteriol. **137:**584–594.