JOURNAL OF BACTERIOLOGY, Mar. 1988, p. 1046–1053
0021-9193/88/031046-08$02.00/0

Vol. 170, No. 3

Structure and Expression of the *Bacillus subtilis sin* Operon

NAND KISHORE GAUR, KETTLY CABANE, AND ISSAR SMITH*

Department of Microbiology, Public Health Research Institute, New York, New York 10016

Received 6 July 1987/Accepted 23 November 1987

The newly identified *sin* gene affects late growth processes in *Bacillus subtilis* when it is overexpressed or inactivated in the chromosome. S1 nuclease mapping of the *sin* gene transcripts in vivo reveals the existence of three transcripts (RNAI, RNAII, and RNAIII). By correlating 5′ ends of *sin* gene transcripts with DNA sequence, we have identified three different promoterlike sequences (P1, P2, and P3) for these transcripts. 3′-End mapping of these transcripts identified three prominent termination sites at the end of the *sin* gene. These termination sites are localized on two hairpin structures previously identified from the DNA sequence. The most abundant transcript, RNAIII, coded only for the *sin* gene, while the polycistronic transcripts RNAII and RNAI coded for the *sin* gene and ORF1 that precedes the *sin* gene. S1 mapping and translational *lacZ* fusion studies indicated that ORF1 and the *sin* gene are regulated differently. ORF1 expression is under developmental control, increasing at the end of vegetative growth, and requires functional *spo0A* and *spo0H* gene products. The *sin* gene is expressed at an almost constant and relatively low level throughout growth and remains largely unaffected by *spo0A* and *spo0H* mutations.

During growth of bacterial cells, the pattern of gene expression changes in response to variations in growth conditions. Bacterial cells adapt to the changing environmental conditions by "turning on" a new set of genes and "turning off" nonessential genes. Multiple genes are often organized into transcriptional units, operons. The organization of genes in an operon helps to achieve coordinate control of gene expression by regulating the amount of transcription from the operon promoter. In a more complex operon, however, discoordinate gene expression is also observed under certain conditions. The α and β, β′ operons of *Escherichia coli* that have genes for core RNA polymerase subunits and ribosomal proteins are examples of such a complex operon (12). The synthesis of RNA polymerase and ribosomes is coordinately regulated upon changes in growth rate, and the regulation is discoordinate during starvation for amino acids. The discoordinate regulation is usually achieved by having internal promoters and terminators within the operon and by regulating expression of genes posttranscriptionally.

Bacillus subtilis undergoes elaborate adaptive changes during transition from vegetative growth to stationary phase in response to nutrient depletion. These changes that accompany and comprise the process of sporulation are production of extracellular enzymes and antibiotics, acquisition of competence, and development of motility (8, 15). Some mutations that block sporulation also prevent appearance of all or a subset of late growth events. Although late growth events are linked to sporulation, they are not required for sporulation. The precise nature of the mechanism that links sporulation with late growth events is not known.

We have previously cloned, characterized, and sequenced a *B. subtilis* DNA fragment which contains a gene, *sin*, that may be involved in controlling late growth events (6). This newly identified gene, when inactivated in the chromosome, displayed pleiotropic phenotypes, which included the loss of ability to develop competence and motility. An overabundance of the *sin* gene product, obtained by gene dosage, was inhibitory to sporulation and production of late growth-specific proteases (neutral and serine proteases) but had no apparent effect on development of competence and motility. The predicted amino acid sequence of the *sin* gene-encoded protein is structurally similar to many bacterial regulatory proteins that have sequence-specific DNA-binding domains, displaying a helix-turn-helix motif. The *sin* gene is therefore likely to be a good candidate for a regulatory gene that may control late growth functions in *B. subtilis*. We describe in this paper the structure of the *sin* gene operon and its regulatory sequences as determined by S1 nuclease mapping of in vivo RNA. The expression of the *sin* gene and ORF1 that precedes the *sin* gene is studied by using in-phase *lacZ* translational fusions.

* Corresponding author.

MATERIALS AND METHODS

Bacterial strains and plasmids. The bacterial strains and plasmids used in this study are listed in Table 1. *B. subtilis* strains were grown on tryptose blood agar base (Difco Laboratories, Detroit, Mich.) or minimal glucose medium and were made competent by the procedure described by Anagnostopoulos and Spizizen (1). *E. coli* JM109 was grown in LB medium and was made competent by the procedure described by Cohen et al. (3). The sporulation inhibition phenotype was measured by colony morphology and by exposing colonies to chloroform vapor (6). Plasmid DNA isolation and transformation were done as described by Gryczan et al. (7). All bacteriological media were made as previously described (6, 10).

Isolation of *B. subtilis* RNA. RNA was prepared from IS170 cells grown in nutrient sporulation medium (NSM) and harvested at different stages of growth. The frozen cell pellet was used to extract RNA by the method described previously (10), but the DNase reaction was done in the following way to remove the chromosomal DNA. The dried RNA pellet suspended in 100 μl of TE buffer (10 mM Tris hydrochloride, pH 8.0, 1 mM EDTA) was treated with 30 U of DNase (Promega Biotec) in the presence of 400 U of RNasin (Promega, Biotec) at 37°C for 30 min, after which RNA was phenol extracted twice with TE-saturated phenol and ethanol precipitated with 25 μl of 3 M ammonium acetate–250 μl of ethanol. The DNA-free RNA pellet was washed with acetone, dried, and suspended in 100 μl of TE.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant genotype or phenotype	Source
B. subtilis		
IS75	*metB5 hisA1 leuA8*	I. Smith
IS170	*metB5 hisA1 recE4*	I. Smith
IS423	*metB5 hisA1 leuA8*(::pIS135)	This work
IS424	*metB5 hisA1 leuA8*(::pIS142)	This work
IS434	*trpC2 pheA1 spo0HΔHind*(::pIS135)	This work
IS435	*trpC2 pheA1 spo0HΔHind*(::pIS142)	This work
IS444	*trpC2 pheA1 spo0AΔ204*(::pIS135)	This work
IS445	*trpC2 pheA1 spo0AΔ204*(::pIS142)	This work
IS457	*trpC2 pheA1*(::pIS135)	This work
IS458	*trpC2 pheA1*(::pIS142)	This work
Plasmid		
pIS74	Cm^r Sin^+	6
pIS112	Cm^r Amp^r	I. Smith
pIS135	Cm^r Amp^r Lac^+	This work
pIS142	Cm^r Amp^r Lac^+	This work

The quality of RNA was determined on 1.5% agarose gel, and the quantity was measured spectrophotometrically.

DNA end labeling. The restriction fragments obtained from pIS74 were 5'-end labeled by removing terminal phosphate with alkaline phosphatase and labeling with [γ-^{32}P]ATP and polynucleotide kinase (10). *Sfa*NI and *Mbo*I end-labeled fragments were restricted with either *Xba*I or *Hin*dIII, and then the fragments were separated on 5% polyacrylamide gel. The appropriate bands were cut out from the gel and isolated by electroelution. Similarly, *Mae*I end-labeled fragments were digested with *Eco*RV, and the appropriate band was isolated.

The 3'-end-labeled probe, also from pIS74, was prepared by using Klenow fragment DNA polymerase (Boehringer Mannheim Biochemicals) and [α-^{32}P]ATP. The *Sfa*NI site, which was used for 3'-end labeling, when restricted, generates a 5' overhang, and from the sequence we determined that the filled in 3' end would have two terminal A nucleotides. The "fill-in" labeling was done with the A^0 mix of Sanger's sequencing method (19). The labeled fragments were digested with *Hin*dIII, and the appropriate band was isolated from the 5% polyacrylamide gel.

S1 nuclease mapping. The labeled probe (10,000 to 20,000 cpm) was mixed with RNA (30 or 150 μg), dried, and suspended in 10 μl of hybridization buffer {80% deionized formamide, 1 mM EDTA, 400 mM NaCl, 40 mM PIPES [piperazine-N,N'-bis (2-ethanesulfonic acid)], pH 6.4}. The mixture was incubated at 80°C for 10 min and then immediately transferred to a water bath set at the hybridization temperature for 3 h. The hybridization temperature was empirically determined to be optimal for the formation of DNA-RNA hybrids. The hybridization temperature, 48°C, was used for all probes except *Mae*I-*Eco*RV and *Mbo*I-*Xba*I probes, for which it was 41°C. The low hybridization temperature was required to see small protected DNA bands. After hybridization, 300 μl of ice-cold S1 buffer (0.28 M NaCl, 30 mM sodium acetate, pH 4.4, 4.5 mM zinc acetate, 20 μg of sonicated calf thymus DNA per ml) and 500 U of S1 nuclease (Boehringer Mannheim) per ml were added and incubated at 37°C for 30 min. The reaction was stopped by adding 100 μl of stop mixture (2.5 M ammonium acetate, 50 mM EDTA). A 1-μl portion of carrier RNA (20 mg/ml) was added and ethanol precipitated. The dried pellet was suspended in 10 μg of loading dye. The samples were electrophoresed on a 6 or 8% polyacrylamide gel containing 7.5 M urea and visualized by autoradiography.

G+A DNA sequencing reaction on the probe was done as described by Maxam and Gilbert (13).

Construction of *lacZ* fusion plasmids pIS135 and pIS142. Plasmid pIS135 (detailed structure shown in Fig. 1) was constructed by cloning an approximately 1-kilobase *Mbo*I fragment of pIS74 (6) containing some vector sequences and all of the regulatory sequences (P1 and P2 and the ribosome-binding site) and 11 codons of the ORF1 gene, terminating at a *Bam*HI site of pIS112 (M. Lewandoski and I. Smith, manuscript in preparation). pIS112 is a derivative of pSK10Δ6 (21) and has a chloramphenicol resistance determinant that can be selected in *B. subtilis*. pIS74 was restricted with *Mbo*I, and the DNA fragments were ligated to *Bam*HI-digested pIS112. The ligated mix was transformed to *E. coli* JM109 selecting for ampicillin resistance (Amp^r) on LB plates containing 100 μg of ampicillin per ml and 5-bromo-4-chloro-3-indolyl-D-galactopyranoside. Plasmid DNA was isolated from transformants which produced blue color on LB–5-bromo-4-chloro-3-indolyl-D-galactopyranoside plates. The presence of plasmids with the proper insert DNA was checked by restricting rapid plasmid preparations with *Eco*RV, which has a unique restriction site in pIS74 in the ORF1-containing insert. From the sequence, we predicted that in-phase cloning would regenerate the *Bam*HI

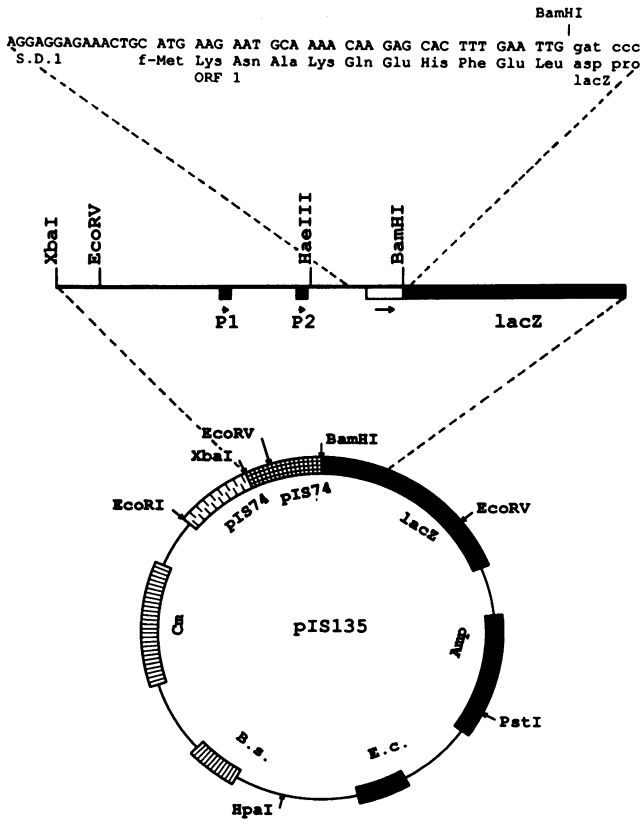

FIG. 1. Structure of pIS135, an ORF1-*lacZ* plasmid. Plasmid construction is described in Materials and Methods. The known regulatory regions (P1, P2, and S.D.1 with 11 N-terminal amino acids of ORF1) which would direct the synthesis of β-galactosidase are shown. The *Eco*RI-*Xba*I fragment upstream of the *Xba*I-*Bam*HI DNA with P1, P2, and S.D.1 comes from the plasmid vector. B.s., *B. subtilis*; E.c., *E. coli*.

site (Fig. 1). The digestion of the plasmid with *Bam*HI showed that the site is regenerated and therefore the ORF1-*lacZ* fusion is in phase. The orientation of the insert was determined by checking the asymmetric *Eco*RV site of the insert relative to flanking *Bam*HI and *Eco*RI sites. The correctness of the ligation and the absence of any mutations in ORF1 have been confirmed by DNA sequencing (data not shown).

pIS142 is a *lacZ* fusion plasmid of the *sin* gene, and its detailed structure is shown in Fig. 2. This plasmid was constructed by using a *Mae*I fragment of pIS74 which includes the proximal part of the ORF1-*sin* gene from *Xba*I (*Mae*I recognizes a subset of *Xba*I sequences) up to a *Mae*I site which is at the 20th codon of the ORF2-*sin* gene. This DNA fragment, after *Mae*I digestion, was purified, filled in with Klenow DNA polymerase, and ligated to *Sma*I-restricted pIS112. The ligated mix was transformed to *E. coli* JM109, and transformants were selected on LB–5-bromo-4-chloro-3-indolyl-D-galactopyranoside plates containing 100 μg of ampicillin per ml. Blue colonies were selected to isolate plasmid DNAs. Those were tested for the *Eco*RV restriction site, which, as above, indicated the presence of the insert. The orientation of the insert was determined by checking the asymmetric *Eco*RV site of the insert relative to flanking *Bam*HI and *Eco*RI sites. The plasmid with the proper insert orientation was designated pIS142. DNA sequencing of pIS142 showed the expected sequence at the fusion site and that upstream sequences were unaltered.

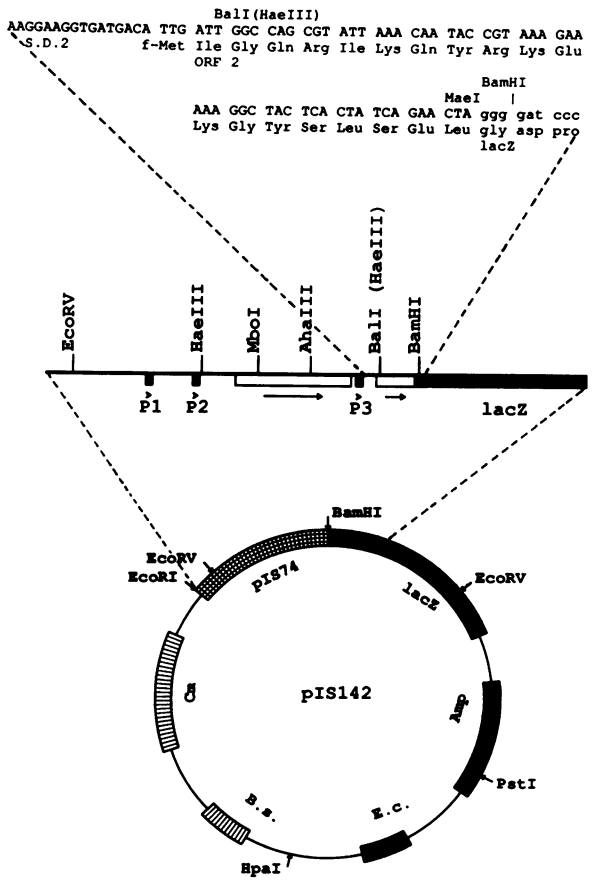

FIG. 2. Structure of pIS142, a *sin-lacZ* fusion plasmid. Construction of the plasmid is described in Materials and Methods. The structure of *sin-lacZ* fusion is shown with known regulatory regions which would direct the synthesis of β-galactosidase. The regulatory regions of interest are *sin* gene promoters (P1, P2, and P3), upstream ORF1, and the *sin* gene translational signal with its 20 N-terminal amino acids. B.s., *B. subtilis*; E.c., *E. coli*.

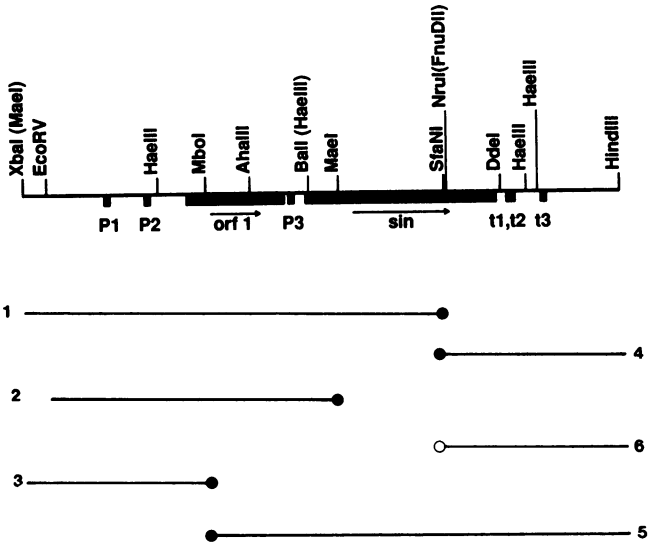

FIG. 3. Probes used for S1 nuclease mapping of *sin* gene transcripts. The pIS74 insert containing ORF1 and the *sin* gene is shown with appropriate restriction sites. Various probes used for S1 nuclease mapping are shown below the pIS74 insert. Filled circles represent labeled 5′ ends. Probe with labeled 3′ end is indicated with the open circle. Unmarked end of the probe is the site of second restriction enzyme.

β-Galactosidase assay. The β-galactosidase activity was determined and activity is expressed in Miller units, as described by Miller (14).

RESULTS

5′-End mapping and regulation of *sin* gene transcripts. The DNA sequence of the fragment containing the *sin* gene showed the existence of five ORFs in both DNA strands (6). Various deletions, mutations, and subcloning analysis suggested that ORF2 encoded the *sin* gene. The DNA sequence also revealed that the *sin* gene was immediately followed by two hairpin structures which resembled *E. coli rho*-independent terminators. Quantitative S1 nuclease mapping experiments were performed to identify the 5′ ends of mRNAs encoding the *sin* gene. IS170 was grown in sporulation medium, and total RNA was extracted from exponentially growing cells (40 to 50 Klett units) and the T_2 stage of sporulation (2 h into sporulation, with T_0 defined as the beginning of the stationary phase of growth). Two levels of RNA (30 and 150 μg) were hybridized with constant and saturating amounts of different probes (Fig. 3) and treated with S1 nuclease (see Materials and Methods). The S1 nuclease-resistant DNA fragments were analyzed on denaturating 6 or 8% polyacrylamide gels. The probe alone with or without S1 treatment was always run as a control.

The 733-base-pair (bp) *Xba*I-*Sfa*N1 probe (probe 1) 5′-end labeled at the *Sfa*NI site was protected by two exponential-phase transcripts, RNAI and RNAIII (Fig. 4). The amount of protection was proportional to the level of total RNA (lanes 1 and 2). The relative amount of RNAIII is at least 10-fold higher than that of RNAI. During sporulation, an additional transcript, RNAII, protected the probe, and the protection

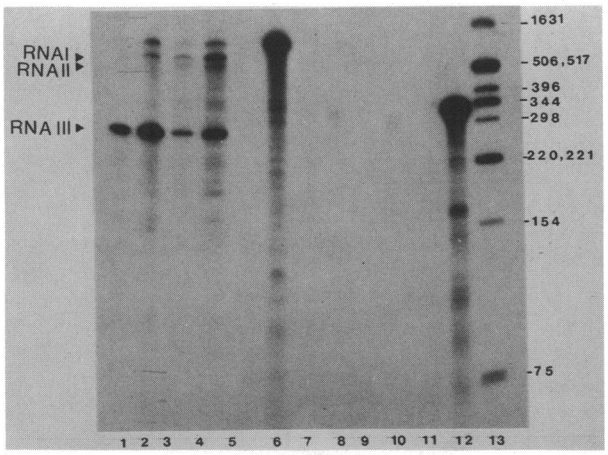

FIG. 4. S1 nuclease mapping of transcripts from the *sin* gene region of the *B. subtilis* chromosome. Total RNA was isolated at exponential phase and a sporulation stage (T_2) from IS170 grown in sporulation medium (NSM). The 733-bp *Xba*I-*Sfa*NI probe that was 5'-end labeled at the *Sfa*NI site was hybridized with two levels of total RNA, S1 nuclease treated, and electrophoresed on an 8% polyacrylamide gel under denaturing conditions (lanes 1 to 6). Similarly, to map any RNA for the other DNA strand, an *Sfa*NI-*Hin*dIII probe (307 bp) that was 5'-end labeled at the *Sfa*NI site was subjected to S1 nuclease mapping (lanes 7 to 12). The samples were run along with 5'-end-labeled *Hin*fI fragments of pBR322, and the sizes are shown in base pairs on the right (lane 13). Lanes 1 through 6, S1 nuclease-protected fragments of *Xba*I-*Sfa*NI probe, hybridized with 30 and 150 μg of exponential-phase RNA (lanes 1 and 2, respectively) and 30 and 150 μg of sporulation-phase (T_2) RNA (lanes 3 and 4), respectively. Controls, Probe with no RNA but treated with S1 (lane 5) or probe alone and not treated with S1 (lane 6). Lanes 7 through 12, S1 nuclease mapping with *Sfa*NI-*Hin*dIII probe, hybridized with 30 and 150 μg of exponential-phase RNA (lanes 7 and 8, respectively) and 30 and 150 μg of sporulation-phase RNA (lanes 9 and 10, respectively). Controls, Probe with no RNA but treated with S1 (lane 11) or probe alone and not treated with S1 (lane 12). The band migrating more slowly than RNAI in lanes 2 and 4 corresponds to a full-length probe not completely digested with S1 nuclease.

increased with increasing amounts of RNA (lanes 3 and 4). Comparison of RNA levels during growth and sporulation suggested that RNAI was increased during sporulation, whereas RNAIII was more abundant during vegetative growth (cf. lanes 1 and 2 with lanes 3 and 4, respectively). RNAII was completely absent during vegetative growth. To determine the precise time of changes in the RNA levels, quantitative S1 nuclease mapping was done with total RNA made at stages T_{-2} T_0, T_2, and T_4 (results not shown). At stage T_0 (beginning of sporulation), the RNAI level increased and the level of RNAIII decreased. RNAII began to appear at stage T_2 and was most abundant at that stage. Approximate sizes of the nuclease-protected DNA fragments were determined relative to labeled single-stranded *Hin*fI-digested pBR322 DNA. These would determine the site of transcription initiation and were calculated to be approximately 580, 500, and 260 bp upstream from the *Sfa*NI site for RNAI, RNAII, and RNAIII, respectively. The band migrating more slowly than RNAI corresponds to the DNA probe not completely digested with S1 nuclease.

To determine more precisely the transcription initiation site, two shorter 5'-end-labeled probes were used. The 549-bp *Eco*RV-*Mae*I probe (probe 2) 5'-end labeled at the *Mae*I site was used mainly to map the 5' end of RNAIII on sequencing gel. As was expected, the probe was protected by all three RNAs, and the protection increased with increasing amounts of total RNA (results not shown). To map the 5' end of RNAIII, S1 nuclease-protected fragments were run along with the G+A Maxam-Gilbert sequencing reaction performed on the same probe (Fig. 5A). The position of the RNAIII-protected band was aligned with the DNA sequence. It corresponded to nucleotide A at position 474 on the published sequence of the *sin* gene (6). It is interesting to note that this position is the beginning of the putative ribosome-binding site of the *sin* gene. It suggested that RNAIII would have no leader sequence 5' upstream to the *sin* gene.

Similarly, the 319-bp *Xba*I-*Mbo*I probe (probe 3) 5'-end labeled at the *Mbo*I site was used to locate the 5' ends of RNAI and RNAII. This probe was protected by RNAI and RNAII but not by RNAIII, and the protection was proportional to the amount of total RNA (data not shown). The S1 nuclease-protected fragments were run beside the G+A

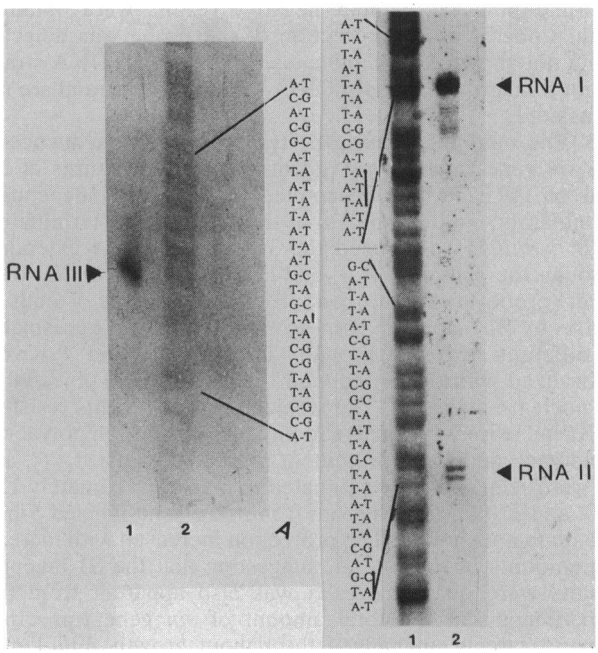

FIG. 5. High-resolution S1 nuclease mapping of 5' ends of *sin* gene transcripts. (A) the 549-bp *Eco*RV-*Mae*I probe 5'-end labeled at the *Mae*I site was subjected to S1 nuclease mapping with 30 and 150 μg of exponential- and sporulation-phase RNA. The samples were analyzed on a low-resolution gel before one of them was run on a 6% polyacrylamide sequencing gel. Lane 1, S1 nuclease-resistant fragment protected by RNAIII, using 30 μg of exponential-phase RNA; lane 2, G+A chemical cleavage ladder of the probe. The DNA sequence of the relevant region is displayed on the right, and the nucleotide corresponding to the 5' end of RNAIII is marked with a bar. The bands on the sequencing gel are light and do not photograph well, but are easily discernible in the original autoradiograph. (B) The 319-bp *Xba*I-*Mbo*I probe 5'-end labeled at the *Mbo*I site was used to map precisely the 5' ends of RNAI and RNAII by S1 nuclease mapping, as described previously. After analysis on low-resolution gel, a sample of S1 nuclease-protected DNA fragments hybridized with 150 μg of sporulation-phase (T_2) RNA was run on 6% polyacrylamide sequencing gel (lane 2). The G+A chemical cleavage ladder of the probe was also run (lane 1). The DNA sequence of the relevant region is displayed on the left. The sequence corresponding to the 5' end of RNAI and RNAII is marked with a vertical line.

Maxam-Gilbert sequence ladder of the same probe (Fig. 5B). The gel position of RNAII- and RNAI-protected DNA fragments was aligned with the DNA sequence to map the precise location of 5′ ends. RNAII-protected fragments corresponded to nucleotides C–A at positions 230 and 231 on the *sin* gene sequence. RNAI-protected fragments corresponded to nucleotide sequence ATA at positions 161 to 163. Multiple protected DNA bands are usually observed in S1 mapping experiments, and the most likely transcription initiation site is considered to be the one corresponding to the largest protected DNA fragment.

The DNA sequence of the *sin* gene-carrying fragment revealed two ORFs in the other DNA strand which, however, had no complete translational signals (6). We decided to determine the existence of any RNA for these ORFs by S1 mapping. Total RNA isolated from exponential and sporulation (T_2) stages was hybridized to an *Sfa*NI-*Hin*dIII probe (probe 4) 5′-end labeled at the *Sfa*NI site and an *Mbo*I-*Hin*dIII probe (probe 5) 5′ labeled at the *Mbo*I site. As with probe 4 (Fig. 4, lanes 7 to 12), no protection was observed at 30 or 150 μg of total RNA prepared from the exponential or sporulation stage. Similar negative results were obtained with probe 5, as no protection of the probe was detected (data not shown). This indicates that the other DNA strand is not transcribed and two ORFs in that DNA strand are not translated.

3′-End mapping of *sin* gene transcripts. The sequence of the *sin* gene revealed two stem-and-loop structures at the end of ORF2 which resembled *E. coli rho*-independent terminators. To determine that they function as terminators in *B. subtilis*, we mapped the 3′ ends of RNAs extending through the *sin* gene. A 3′-end-labeled *Sfa*NI-*Hin*dIII fragment (probe 6) (308 bp) was used to probe the 3′ ends of RNAs by S1 mapping. Total RNAs, 30 and 150 μg, isolated at different stages of growth (T_{-2}, T_0, T_2, and T_4) were hybridized with the probe, and the hybrid was subjected to S1 nuclease digestion. The labeled DNA fragments resistant to S1 nuclease were analyzed on a low-resolution polyacrylamide gel, and three prominent protected bands, t_1, t_2, and t_3, were seen and were estimated to map approximately 120, 130, and 170 bp downstream from the 3′-end-labeled *Sfa*NI site (data not shown). The protection increased with increasing amounts of total RNA, suggesting that the S1 mapping results were quantitative. It was also apparent from this experiment that the total amount of *sin* gene transcripts remains largely unchanged throughout growth. Full-length protection of the probe was not observed, indicating that all transcripts extending through the *sin* gene terminated at one of these three termination sites. It would suggest that there would be no readthrough transcript available for the translation of ORF3, an ORF distal to the *sin* gene (6). However, we cannot rule out any RNA originating downstream from the *Sfa*NI site which can be transcribed through these terminators.

To precisely map the 3′ ends of RNAs, S1 nuclease-protected fragments were run beside the G+A Maxam-Gilbert sequence ladder of the probe on sequencing gel (Fig. 6). The positions of S1 nuclease-resistant bands were aligned with the corresponding sequence. Figure 7 shows the location of major protected bands displayed on the predicted hairpin structures. It is interesting to note that the first hairpin structure has two major termination sites (t_1 and t_2). It is known in *E. coli rho*-independent terminators that the major RNA termination site is the first T at the base of the stem (18). It is possible that the first hairpin structure exists in two alternate structures, the second one being open up to

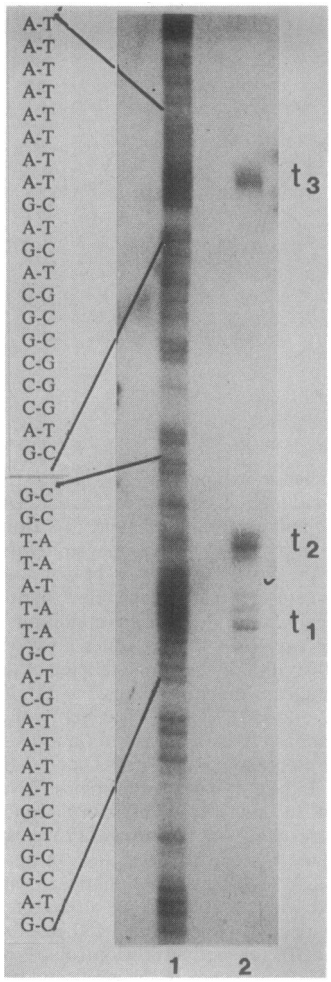

FIG. 6. Termination of *sin* gene transcripts. High-resolution S1 nuclease mapping of 3′ ends of *sin* gene transcripts. S1 nuclease-resistant DNA fragments of a 3′-end-labeled *Sfa*NI-*Hin*dIII probe hybridized with 150 μg of T_2 RNA were run on a 7.5 M urea–6% polyacrylamide gel (lane 2). The G+A chemical cleavage ladder of the probe is shown in lane 1. The DNA sequence of the relevant region taken from the published sequence (6) is displayed on the left.

t_1 at the base of the stem. In fact, there are five A or T nucleotides out of a total of six in the paired region below t_1, which in an equilibrium with paired and unpaired structures would favor the latter. However, the stem above t_1 should be stable as it is rich in G-C base pairs. The only major termination site, t_3, in the second hairpin structure is at the first T, which is preceded by a stem having 7 G-C bp of 11 bp.

Expression of ORF1-*lacZ* and *sin-lacZ* fusions. After learning that the DNA region encoding ORF1 and the *sin* gene is transcribed in vivo, we decided to make translational *lacZ* fusions to demonstrate that they indeed are translated. These fusions were also used to study regulation under various physiological conditions. The in-phase *lacZ* fusion for ORF1 was constructed at the *Mbo*I site generating pIS135 (see Materials and Methods), which would use ORF1 transcriptional and translational signals for expression of β-galactosidase. A similar transcriptional-translational fusion was constructed between *lacZ* and the *sin* gene at the *Mae*I site, creating a plasmid, pIS142 (see Materials and Methods for details). pIS135 and pIS142 have defective

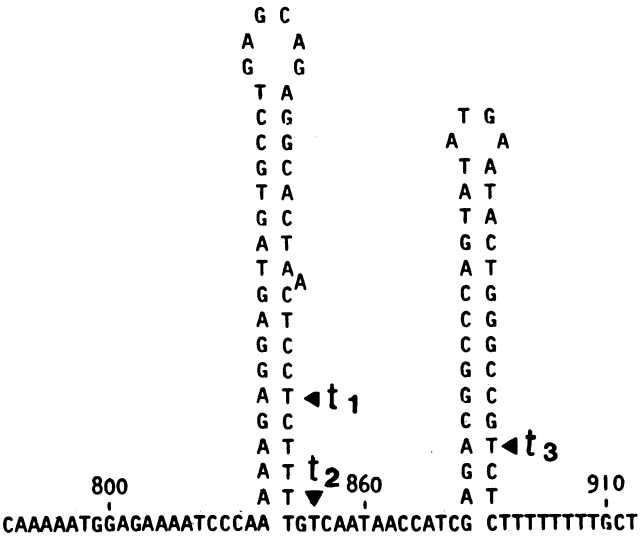

FIG. 7. Structure of two terminators located at the end of the *sin* gene. The most prominent band representing each termination site is indicated with an arrow on the DNA sequence presented as hairpin structures.

replicative function in *B. subtilis* and are thus integrative plasmids. A wild-type strain, IS75, was transformed with pIS135 and pIS142 by selecting for plasmid-encoded chloramphenicol resistance. The Cm^r transformants would originate through a Campbell-like integration of the whole plasmid into the chromosome, using the homology of the cloned *sin* gene DNA sequences. The integration of the plasmid into the chromosome would result in the expression of β-galactosidase synthesis under the control of any known or unknown control elements of the *sin* gene which lie upstream from the *lacZ* fusion site. IS423 and IS424 strains were constructed this way by transformation of pIS135 and pIS142, respectively. β-Galactosidase synthesis in IS423 would be under the control of promoters P1 and P2 and translational initiation signals of ORF1. Similarly, in IS424, promoters P1, P2, and P3 would direct the synthesis of β-galactosidase. However, in IS424, ORF2 translational signals would be used for β-galactosidase synthesis. Chromosomal DNAs prepared from IS423 and IS424 were used to transfer the *lacZ* fusions to various genetic backgrounds.

The β-galactosidase activity of ORF1-*lacZ* and the *sin-lacZ* fusions was measured at 0.5-h intervals under various conditions. The β-galactosidase activity of ORF1-*lacZ* fusion in a wild-type Spo^+ strain (IS457) was induced three- to fourfold at the beginning of sporulation (Fig. 8). This induction was prevented by supplementing NSM with 0.5% glucose, which also suppressed sporulation. The effects of two *spo0* mutations, *spo0A* and *spo0H*, were also studied on the expression of ORF1-*lacZ* fusion. Figure 8 showed that the β-galactosidase activity was dramatically reduced in the *spo0H* mutant. The activity was also very low in the *spo0A* mutant. The β-galactosidase activity of the *sin-lacZ* fusion in the wild-type strain (IS458) remained relatively unchanged during growth and stationary phases when grown in NSM with or without glucose (Fig. 9). Contrary to ORF1-*lacZ* fusions, the *sin-lacZ* fusion activity remained largely unaffected by *spo0A* or *spo0H* mutations. While *sin-lacZ* activity remains low throughout growth and stationary phases, it is much higher than endogenous β-galactosidase activity (in the absence of any integrated *lacZ* fusions), which reaches 1 to 2 Miller units during stationary phase and is completely repressed by glucose (4).

Structural features of *sin* gene promoters. Three RNAs (RNAI, RNAII, and RNAIII), as revealed by S1 nuclease mapping, would be transcribed through the *sin* gene. The time of appearance and relative abundance of these RNAs suggested that they might be transcribed by a different type of promoters or were regulated differently or both. If the 5′ ends of these RNAs in vivo were unprocessed and represented the actual transcriptional initiation site, then the examination of DNA sequence in the −10 and −35 regions should allow us to tentatively assign the type of promoters for these RNAs. P3, the promoter for RNAIII, has a −10 region identical to σ^{43} promoters (11) and has two of six bases identical in the −35 region (Table 2). The promoter for RNAII (P2) is tentatively assigned to be a σ^{29} promoter for the following reasons: (i) the time of appearance of RNAII is the same as of promoters utilized by σ^{29} (17); (ii) the DNA sequences in −10 and −35 regions are similar to consensus σ^{29} promoter sequences (Table 2). P1, the promoter for RNAI, is induced at the beginning of sporulation and its activity is severely impaired in the *spo0H* and *spo0A* mutants. As the *spo0H* protein is now considered to be a newly discovered sigma factor, σ^{30} (4a) which is part of a holoenzyme that reads the *spoVG* promoter (2), it is possible that P1 is recognized by σ^{30}. The DNA sequences of P1 are compared with *citG* promoter P2, which also requires *spo0H* for its in vivo transcription (I. Feavers and A. Moir, personal communication). It is interesting to note that 6 bp upstream

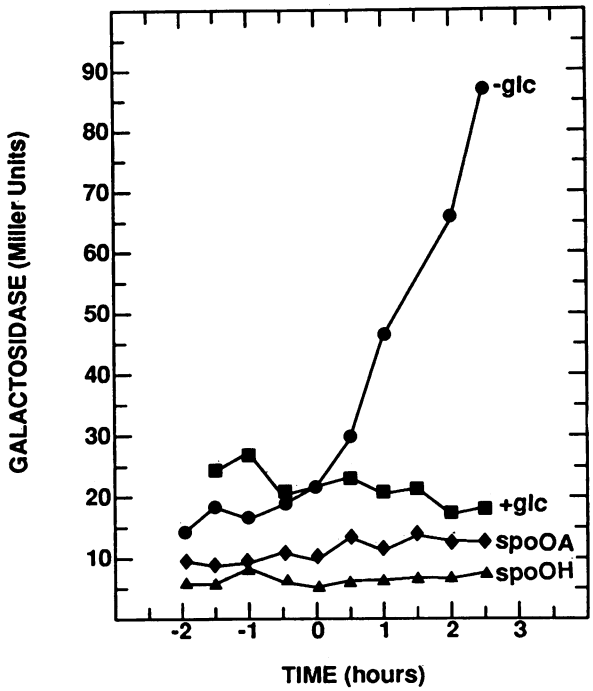

FIG. 8. Regulation of ORF1-*lacZ* fusion integrated in the chromosome and measured as β-galactosidase activity. The ordinate represents β-galactosidase activity expressed in Miller units (14). The abscissa represents growth time in hours, and T_0 is the transition point between vegetative growth and stationary phase. Symbols: ●, IS457 (Spo^+) grown in NSM containing 5 μg of chloramphenicol per ml; ■, IS457 (Spo^+) grown in NSM containing 0.5% glucose and 5 μg of chloramphenicol per ml; ◆, IS444 (*spo0A*) grown in NSM containing 5 μg of chloramphenicol per ml; ▲, IS434 (*spo0H*) grown in NSM containing 5 μg of chloramphenicol per ml.

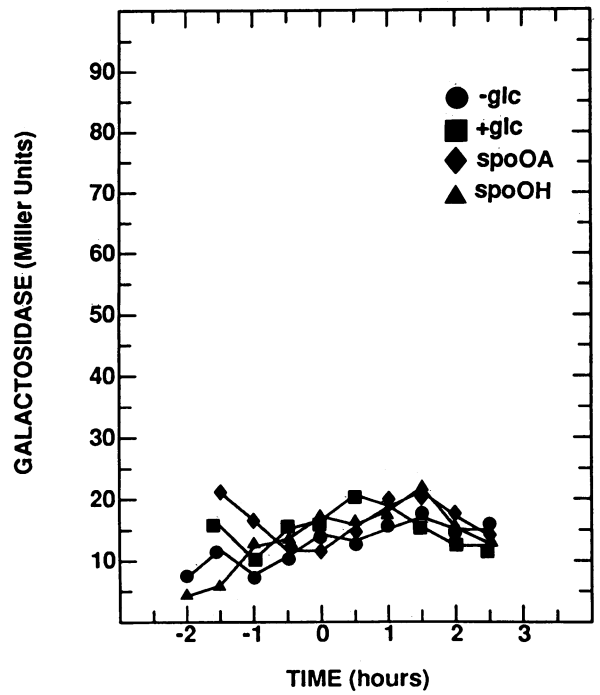

FIG. 9. Regulation of chromosome integrated *sin-lacZ* fusion measured as β-galactosidase activity. The ordinate and abscissa are as in the legend to Fig. 8, with the same scale for comparison. Symbols: ●, IS458 (Spo$^+$) grown in NSM containing 5 μg of chloramphenicol per ml; ■, IS458 (Spo$^+$) grown in NSM containing 0.5% glucose and 5 μg of chloramphenicol per ml; ◆, IS445 (*spo0A*) grown in NSM containing 5 μg of chloramphenicol per ml; ▲, IS435 (*spo0H*) grown in NSM containing 5 μg of chloramphenicol per ml.

from the transcription start site it has 7 of 8 bp identical to that of *citG* promoter P2 (Table 2). There is also identity of 7 of 8 bp in the −35 region when *citG* P2 and *sin* P1 are compared. However, there is a TATAAT sequence at −7 to −12, so it is also possible that this promoter is transcribed by σ^{43} polymerase.

DISCUSSION

The *sin* gene was originally identified as a gene which had an inhibitory effect on *B. subtilis* sporulation when cloned in multicopy plasmids (6). The inhibition activity was localized to a DNA sequence containing an open reading frame, ORF2, now named *sin*, preceded by a regulatory region. DNA sequence analysis showed that the *sin* gene ORF2 had a DNA-binding domain similar to that of many procaryotic regulatory proteins. ORF1, which preceded *sin* ORF2, was not required for multicopy inhibition of sporulation and protease production. The inactivation of the *sin* gene and ORF1 in the chromosome caused an overproduction of α-amylase and protease activity and the loss of competence and motility.

Combining our data of S1 nuclease mapping for both 5′ and 3′ ends of in vivo RNA from ORF1 and the *sin* gene regions, we have shown the existence of three transcripts for the *sin* gene. The major transcript, RNAIII, coding only for the *sin* gene, is transcribed from a σ^{43}-like promoter (P3). Two larger transcripts, RNAI and RNAII, initiate at promoters P1 and P2, extend through ORF1 and the *sin* gene regions, and are thus polycistronic. Promoters P1 and P2 may be utilized by alternative RNA polymerases, but the identity of the RNA polymerase forms utilizing these promoters awaits in vitro transcription studies. These three transcripts terminate at two terminators similar to *E. coli rho*-independent terminators which are located at the end of the *sin* gene.

The *lacZ* fusions in ORF1 and the *sin* gene and the existence of transcripts in this region unequivocally demonstrate that they are translated in vivo. It is interesting to note that ORF1-*lacZ* expression is higher than the *sin-lacZ* expression even though the total RNA for the *sin* gene is much higher than the ORF1 RNA. We feel that the low levels of *sin* gene expression are not artifacts due to the construction and integration of the *lacZ* fusions. We have recently shown that the chromosomal *sin* gene is expressed at extremely low levels throughout growth, as evidenced by immunodetection of the *sin* protein (Gaur and Smith, unpublished results). Clearly, *sin-lacZ* expression is regulated posttranscriptionally. It is possible that the *sin* gene has poor translational signals and therefore is expressed poorly. The comparison of putative ribosome-binding sites indicates that indeed the *sin* gene has a relatively poor ribosome-binding site (Fig. 1 and 2), but it is not possible, as yet, to conclude that this difference in ribosome-binding site is enough to account for such a difference in ORF1-*lacZ* and *sin-lacZ* expressions. Additional posttranscriptional controls cannot be completely ruled out. We can, however, rule out the involvement of any countertranscript which might prevent the

TABLE 2. Comparison of promoter DNA sequences[a]

Promoter	−35	Spacer	−10	
	* *		* * * * * *	
σ^{43}	T T G A C A		T A T A A T	
sin P3	C C G T A A A T C C T T T C T G A A T G T G C T A T A A T A T C A C A			+ 1
	* *		* * * * *	
σ^{29}	G A A - A A - T		C A T A T T - T	
sin P2	G A C T T C C A G A G A C T A A T G A A G C A T A C A A T A A G T C A			+ 1
	* * * * * * * * * *			
spoVG	A G G A T T T C A G A A A A A A T C G T		G G A A T T G A T A C A C T A A T G	
citG P2	A G G A T T T T T T G T G T C A T T G G		C G A A T T T A T G A T C T A T T G A	
sin P1	C T C G T T T T T T T T G A G A A A A A A T A C G A T T A T A A T A A A G G T A			+ 1
	* * * * * * * * * *		* * * * * * * * *	

[a] The sequence of a nontranscribed strand is given. The +1 positions of the *sin* gene promoters are indicated. Identical nucleotides are marked with asterisks (*). The consensus promoter sequences of σ^{43} (11) and σ^{29} (17) are shown above the *sin* gene promoters P3 and P2. The *sin* gene promoter P1 is compared with the *citG* promoter P2 (Feavers and Moir, personal communication) and the *spoVG* promoter utilized by σ^{30} in vitro (2).

translation of *sin* gene transcripts as has been observed for some genes (16). We have not detected any transcript from the other DNA strand in vivo by using two different probes.

ORF1-*lacZ* expression is induced at the beginning of sporulation and induction is prevented by high glucose, which also prevents sporulation. It suggests that induction is related to sporulation. The increased level of RNAI at the beginning of sporulation, as has been observed by S1 nuclease mapping, may account for this induced ORF1-*lacZ* expression. RNAII also increases, but this rise occurs at T_2, and this cannot account for the T_0 rise. ORF1-*lacZ* expression is reduced in the *spo0A* strain and more so in a *spo0H* background. It suggested the role of these *spo0* genes during vegetative growth. Perhaps the *spo0H* gene product, being a sigma factor (σ^{30}) (2; Dubnau et al., in preparation) controls the utilization of promoter P1. The observation that *spo0A* is required for expression of a *spo0H-lacZ* fusion may explain the *spo0A*-dependent expression of ORF1-*lacZ* (4; J. Weir et al., unpublished results).

It is interesting to note that *sin-lacZ* fusion expression is relatively unchanged under those conditions that affect ORF1-*lacZ* fusion expression even though ORF1 and the *sin* gene are in the same operon. It is possible to speculate on the nature of this discoordinate regulation of expression of ORF1 and the *sin* gene. The steady-state levels of *sin* gene transcripts may represent the extent of contribution they make for the formation of *sin* gene protein. In that case the most abundant transcript, RNAIII, should contribute maximally to the formation of the *sin* gene protein. The contribution of relatively small amounts of polycistronic transcripts (RNAI and RNAII) for the formation of *sin* gene protein may be minimal. Promoter P3 for RNAIII may not be subjected to developmental regulation, and therefore the expression of *sin* gene is not significantly affected by mutations in *spo0A* and *spo0H* genes. However, as ORF1 is translated with only polycistronic transcripts, its expression would be dependent on promoter P1 and, after T_2, on both promoters P1 and P2.

In *B. subtilis*, many genes with multiple promoters and genes in an operon have been reported previously (5, 9, 10, 20). However, the *sin* gene operon may be the first example in which the expression of an upstream gene (ORF1) is sensitive to high glucose and mutations in *spo0* genes, and the expression of downstream gene (*sin* gene) remains relatively unaffected under these conditions.

While the regulation of ORF1 and the *sin* genes is of interest in our understanding of late growth control, the actual function of the *sin* protein is also significant in this regard. As previously described, overproduction of the *sin* protein inhibits protease production, as well as sporulation, and disruption of the *sin* gene causes a loss of competence (6). Recently, we have shown that the *sin* protein behaves like a DNA-binding protein and in vivo inhibits the transcription of the *aprE* gene (Gaur and Smith, unpublished results). *B. subtilis* strains with a *sin* gene disruption are completely negative with respect to the in vivo transcription of several competence genes (Y. Weinrauch and D. Dubnau, personal communication). This suggests that the *sin* protein is a regulator, both negative and positive, of late growth genes.

ACKNOWLEDGMENTS

We thank Eugenie Dubnau, Mark Lewandoski, and Joyce Weir for helpful discussions and Ed Chang for obtaining chemical cleavage ladders. We also thank Annabel Howard for expert secretarial assistance and James Hagenzieker for photographical assistance.

This work was supported by Public Health Service grant GM-32651 from the National Institute of General Medical Sciences awarded to I.S.

LITERATURE CITED

1. **Anagnostopoulos, C., and J. Spizizen.** 1961. Requirements for transformation in *Bacillus subtilis*. J. Bacteriol. **81**:741–746.
2. **Carter, H. L., III, and C. P. Moran, Jr.** 1986. New RNA polymerase σ factor under *spo0* control in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **83**:9438–9442.
3. **Cohen, S. N., A. C. Y. Chang, H. Boyer, and R. Helling.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc. Nat. Acad. Sci. USA **69**:2110–2114.
4. **Dubnau, E., K. Cabane, and I. Smith.** 1987. Regulation of *spo0H*, an early sporulation gene in bacilli. J. Bacteriol. **169**: 1182–1191.

4a. **Dubnau, E., J. Weir, G. Nair, L. Carter III, C. Moran, Jr., and I. Smith.** 1988. *Bacillus* sporulation gene *spo0H* codes for σ^{30} (σ^{H}). J. Bacteriol. **170**:1054–1062.

5. **Ferrari, F. A., K. Trach, and J. A. Hoch.** 1985. Sequence analysis of the *spo0B* locus reveals a polycistronic transcription unit. J. Bacteriol. **161**:556–562.
6. **Gaur, N. K., E. Dubnau, and I. Smith.** 1986. Characterization of a cloned *Bacillus subtilis* gene that inhibits sporulation in multiple copies. J. Bacteriol. **168**:860–869.
7. **Gryczan, T. J., S. Contente, and D. Dubnau.** 1978. Characterization of *Staphylococcus aureus* plasmids introduced by transformation into *Bacillus subtilis*. J. Bacteriol. **134**:318–329.
8. **Hoch, J. A.** 1976. Genetics of bacterial sporulation. Adv. Genet. **18**:69–99.
9. **Johnson, W. C., C. P. Moran, Jr., and R. Losick.** 1983. Two RNA polymerase sigma factors from *Bacillus subtilis* discriminate between overlapping promoters for a developmentally regulated gene. Nature (London) **302**:800-804.
10. **Lewandoski, M., E. Dubnau, and I. Smith.** 1986. Transcriptional regulation of the *spo0F* gene of *Bacillus subtilis*. J. Bacteriol. **168**:870–877.
11. **Losick, R., and J. Pero.** 1981. Cascades of sigma factors. Cell **25**:582–584.
12. **Maher, D. L., and P. P. Dennis.** 1977. In vivo transcription of *E. coli* genes coding for rRNA, ribosomal proteins and subunits of RNA polymerase: influence of the stringent control system. Mol. Gen. Genet. **155**:203–211.
13. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**: 499–560.
14. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Piggot, P. J., and J. G. Coote.** 1976. Genetic aspects of bacterial endospore formation. Bacteriol. Rev. **40**:908–962.
16. **Pines, O., and M. Inouye.** 1986. Antisense RNA regulation in prokaryotes. Trends Genet. **2**:284–287.
17. **Rather, P. N., R. E. Hay, G. L. Ray, W. G. Haldenwang, and C. P. Moran, Jr.** 1986. Nucleotide sequences that define promoters that are used by *Bacillus subtilis* sigma-29 RNA polymerase. J. Mol. Biol. **192**:557–565.
18. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
19. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
20. **Wang, L. F., and R. H. Doi.** 1986. Nucleotide sequence and organization of *Bacillus subtilis* RNA polymerase major sigma (σ^{43}) operon. Nucleic Acids Res. **14**:4293–4307.
21. **Zuber, P., and R. Losick.** 1983. Use of a *lacZ* fusion to study the role of the *spo0* genes of Bacillus subtilis in developmental regulation. Cell **35**:275–283.