JOURNAL OF BACTERIOLOGY, Nov. 1997, p. 7046–7054 Vol. 179, No. 22
0021-9193/97/$04.00+0


# Analysis of the *Bacillus subtilis* S10 Ribosomal Protein Gene Cluster Identifies Two Promoters That May Be Responsible for Transcription of the Entire 15-Kilobase S10-*spc*-α Cluster

XIAO LI, LASSE LINDAHL, YIZHONG SHA, AND JANICE M. ZENGEL*

*Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, Maryland 21250*

Received 1 July 1997/Accepted 16 September 1997

**We have sequenced a previously uncharacterized region of the *Bacillus subtilis* S10 ribosomal protein gene cluster. The new segment includes genes for S10, L3, L4, L23, L2, S19, L22, S3, and part of L16. These *B. subtilis* genes map in the same order as the genes in the *Escherichia coli* S10 ribosomal protein operon. Two potential promoter sequences were identified, one ~200 bases and the other ~140 bases upstream of the S10 gene. The activities of the two promoters were demonstrated by primer extension analysis, in vitro transcription experiments, and in vivo promoter fusion plasmid studies. In agreement with previous reports, our Northern analysis of exponentially growing cells failed to identify terminators or other active promoters within the S10-*spc*-α region. Our observations suggest that the two S10 promoters reported here are responsible for transcribing a 15-kb-long transcript for all of the genes in the *B. subtilis* S10, *spc*, and α clusters.**

In *Escherichia coli*, ribosomal protein (r-protein) synthesis is regulated at the level of translation by an autogenous feedback mechanism (reviewed in reference 40). In most of the r-protein operons studied, one of the gene products acts as an operon-specific translational repressor when it accumulates in amounts exceeding the available binding sites on nascent ribosomal particles (25, 40). In the S10 operon the translational regulatory r-protein, L4, also regulates transcription by inducing premature termination at a specific site within the leader (20). Even though much is known about the regulation of r-protein synthesis in *E. coli*, information about this process in other bacterial species is still sketchy. For example, regulation of only one r-protein gene in *Bacillus subtilis*, coding for S4, has been characterized (11). This paper addresses the genetic organization and regulation of a cluster of r-proteins in *B. subtilis*. Since transcription units and regulatory mechanisms have not yet been characterized for *B. subtilis*, we use the terms S10, *spc*, and α "gene clusters" to refer to the groups of *B. subtilis* genes that correspond to the S10, *spc*, and α operons in *E. coli.*

A major r-protein gene cluster is located at 12° on the *B. subtilis* chromosome. Previous work in other laboratories has demonstrated that this region harbors most of the genes found in the *E. coli* L11, β, S10, *spc*, and α operons, as well as other genes such as the gene for initiation factor IF1 (1, 13). DNA sequencing is complete for the region analogous to the *E. coli spc* and α operons (1, 15, 34). Of the *B. subtilis* genes corresponding to the 11-gene S10 operon of *E. coli*, only the sequences of the L29, S17, and L16 (partial) genes have been reported (15, 34). All of these identified r-protein genes (the S4 gene maps elsewhere; see below) are arranged in the same order in the two genomes.

Even though the overall gene organization of this cluster is similar in gram-positive and gram-negative organisms, the transcriptional organization seems to be different. A terminator structure with dyad symmetry has been identified at the 3′ end of the S10 r-protein operon in *E. coli* (28) and shown to terminate about 80% of the transcripts initiated from the S10 promoter (21). No such terminator structure has been identified in the corresponding region of the *B. subtilis* sequence (15). In addition, no major promoters have been found in the sequence from L16 (S10 cluster) to the end of the α cluster (34). Several weak promoters were identified in regions different from the locations of the *E. coli* promoters (34). These earlier results led to the suggestion that the entire S10-*spc*-α region might constitute a single transcription unit (15, 34).

The regulatory mechanisms governing the expression of the major r-protein gene cluster also seem to be different in the two bacterial species. In *E. coli*, translation of the α operon is autogenously regulated by the product of the third gene in the operon, S4 (17). However, in *B. subtilis* the S4 gene is located in a single gene operon at 263° on the chromosome (14), and the protein autogenously regulates only its own expression (10, 11). The regulator of the remaining genes in the *B. subtilis* α cluster remains unknown. Similarly, in *E. coli*, translation of the *spc* operon mRNA is regulated by binding of the S8 protein to a bulged hairpin, located near the beginning of the L5 gene, which has substantial structural homology with the S8 binding site on 16S rRNA (4, 9). In *B. subtilis*, the sequence upstream of the L5 gene is not compatible with the formation of a similar target for the r-protein S8 (15). The questions arise whether the S10 gene, *spc*, and α clusters form a single transcription unit in *B. subtilis* and, if so, whether a single r-protein (for example, L4, the regulatory protein of the S10 operon in *E. coli* [22]) regulates expression of the entire S10-*spc*-α region in *B. subtilis*.

To address these questions, we have sequenced the previously uncharacterized part of the S10 cluster from *B. subtilis*. We found that the gene organization of this region in *B. subtilis* is identical to that of the S10 operon of *E. coli*. Furthermore, we identified two active promoters upstream of the S10 structural gene. Our experiments yielded no evidence of transcription from promoters internal to the S10-*spc*-α cluster. Hence, our results suggest that the genes in the S10-*spc*-α region might constitute a single transcription unit.

* Corresponding author. Mailing address: Department of Biological Sciences, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250. Phone: (410) 455-2876. Fax: (410) 455-3875. E-mail: zengel@umbc.edu.

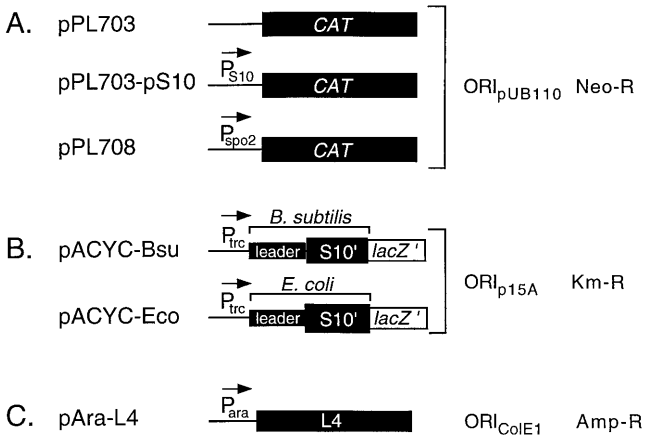


FIG. 1. (A) Plasmid pPL703 and derivatives (reference 7 and this report) used for analyzing the promoter activity of the sequence upstream of the S10 gene. These plasmids contain the origin of replication of the *B. subtilis* plasmid pUB110. The *B. subtilis* DNA fragment cloned in pPL703-pS10 contained nucleotides 1 to 206 of the sequence shown in Fig. 2. (B) Plasmids used to analyze the effect of *E. coli* L4 on expression of the *B. subtilis* leader-S10–*lacZ* fusion. These plasmids all contain the *E. coli* p15A origin of replication. The *B. subtilis* DNA cloned in pACYC-Bsu contains nucleotides 286 to 480 of the sequence shown in Fig. 2. (C) Plasmid used for oversynthesis of *E. coli* L4, containing the ColE1 origin of replication.

## MATERIALS AND METHODS

**Strains, plasmids, and growth conditions.** We used *E. coli* K-12 strain LL308 (22) and *B. subtilis* SG38 (from J. R. McCormick and R. Losick). Cells were grown at 37°C in Luria-Bertani (LB) liquid or on TBAB (Difco) plates.

pPL703 is a *B. subtilis* plasmid that contains a gene for neomycin resistance and a promoterless chloramphenicol acetyltransferase (CAT) gene (7). The region of DNA containing potential promoters of the *B. subtilis* S10 operon was cloned in front of the CAT gene to create plasmid pPL703-pS10 (Fig. 1A). Plasmid pPL708 contains a promoter from the SPO2 phage of *B. subtilis* in front of the CAT gene of pPL703 (7) (Fig. 1A). Plasmid pACYC-*lacZ* is a derivative of *E. coli* plasmid pACYC-CΔ1 (19), containing the *trc* promoter of pTrc99A (Pharmacia) and a *lacZ* gene lacking its Shine-Dalgarno sequence and the first 8 codons. Plasmid pACYC-Bsu was constructed from pACYC-*lacZ* by inserting the S10 leader and proximal 56 codons of the S10 gene from *B. subtilis* between the *trc* promoter and *lacZ* gene, with the S10 gene fused in frame to *lacZ* (Fig. 1B). The analogous construct containing the *E. coli* S10 leader and proximal S10 gene was also constructed and is called pACYC-Eco (Fig. 1B). The pAra-L4 plasmid, another *E. coli* plasmid, contains the L4 gene of *E. coli* under control of the arabinose-inducible $P_{ara}$ promoter (3) (Fig. 1C).

**Sequencing of the *B. subtilis* S10 operon.** To sequence the genes between those for L4 and L16, a 3-kb fragment was amplified from *B. subtilis* SG38 chromosomal DNA by PCR using an upstream primer that hybridizes to a region of the L4 gene and a downstream primer that hybridizes to the L16 gene of *B. subtilis*. The sequence for the L4 gene oligonucleotide was AAACCGTGGCGTCAAA AAGGAACT, encoding the amino acid sequence KPWRQKGT, which is highly conserved in eubacterial L4 proteins (19, 41); the DNA sequence contains codons that we determined are most frequently used in *B. subtilis* r-protein genes. The L16 gene oligonucleotide was TTTGACGGTTAGTAATCCAAGAAG, from the published sequence of the *B. subtilis* L16 gene (15). The PCR fragments from two reactions were each purified from an agarose gel (Gene Clean kit; Bio 101) and sequenced (29) in both orientations.

To sequence genes upstream of the L4 gene, SG38 chromosomal DNA was digested with *Eco*RI and analyzed by Southern blotting (23). The DNA probe was a 0.6-kb fragment from the *Bacillus stearothermophilus* L4 gene (16) labeled by random priming. A band approximately 3 kb in size was found to hybridize to the probe. DNA fragments of that size were purified from an agarose gel with GeneClean and ligated to an *Eco*RI-*Pst*I fragment of plasmid pLL226 (37). A 1.5-kb DNA fragment containing the genes for S10, L3, and the unknown part of L4 of *B. subtilis* was amplified by PCR using primers which hybridize to pLL226 and the known part of the L4 gene. The fragment was sequenced as described above.

**Labeling and gel electrophoresis of proteins.** *E. coli* cells were grown at 37°C in AB minimal medium (6) supplemented with 0.5% glycerol, 1 μg of thiamine per ml, and the appropriate antibiotics. The synthesis of the S10–β-galactosidase fusion protein was induced by addition of IPTG (isopropyl-β-D-thiogalactopyranoside) (to 1 mM). Immediately before and 10 min after addition of IPTG, aliquots of the culture were pulse-labeled for 2 min with 40 μCi of [$^{35}$S]methionine per ml and lysed at 95°C in sodium dodecyl sulfate sample buffer (18). Synthesis of L4 protein was induced by the addition of arabinose (to 0.5%). Ten minutes later aliquots were pulse-labeled and harvested as described above. The total extracts were fractionated on a sodium dodecyl sulfate–7.5% polyacrylamide gel (18) and analyzed by autoradiography.

**Primer extension assay.** Total RNA was isolated by the hot-phenol method (5) from *B. subtilis* SG38 grown to mid-log phase. Primers hybridizing to either the S10 leader (nucleotides 263 to 278) (Fig. 2) or the S10 structural gene (nucleotides 374 to 390) (Fig. 2) were end labeled with [γ-$^{32}$P]ATP by T4 kinase. Primer ($10^6$ cpm) was mixed with 10 μg of *B. subtilis* RNA in 50 mM Tris-Cl (pH 8.0)–100 mM KCl buffer, incubated at 90°C for 1 min, and slowly cooled to room temperature. The RNA-primer mixture was mixed with 5× reverse transcriptase buffer (Bethesda Research Laboratories [BRL]), 10 mM dithiothreitol, 0.1 mM deoxyribonucleoside triphosphates, 25 U of RNasin (from Promega), and 200 U of Moloney murine leukemia virus Superscript-II reverse transcriptase (BRL) and incubated at room temperature for 5 min and then at 52°C for 1 h. Reactions were terminated by the addition of 5 μl of stop mixture, and reaction products were analyzed by electrophoresis on an 8% polyacrylamide gel containing 8 M urea.

**Northern analysis.** Total RNA from SG38 grown to either mid-log phase or stationary phase was electrophoresed on a 0.8 or 1% formaldehyde-agarose gel and transferred to a nylon membrane by capillary blotting. Northern blot analysis was carried out as described previously (23). Probes from various genes in the *B. subtilis* S10-*spc*-α region were first amplified from genomic DNA by PCR and then randomly labeled with [α-$^{32}$P]dATP (random-labeling kit from Promega).

**In vitro transcription reactions.** Standard 10-μl transcription reaction mixtures contained 20 mM Tris-acetate (pH 7.9), 4 mM magnesium acetate, 0.1 mM EDTA, 100 mM potassium glutamate, 20 nM *E. coli* RNA polymerase, and 20 nM PCR DNA fragment. The PCR fragment of the *B. subtilis* S10 leader was amplified from chromosomal DNA by using two oligonucleotides: the upstream oligonucleotide 0218, which anneals to a region 5′ of the two S10 promoters, has the sequence CCCAAGCTTCAGCTCACACCCCGCATA, and the downstream oligonucleotide 0224, which anneals to a region within the S10 structural gene, has the sequence GGGGTACCGCCAGAAGGATTGTGTAAA. The PCR fragment of the *E. coli* S10 leader was amplified by using an upstream oligonucleotide (GAGAGCGTGTCAAAAATGCA), which anneals to a region about 100 nucleotides 5′ of the S10 promoter, and a downstream oligonucleotide (GCTCTTTGCGTGTCGGCAGC), which anneals to a region within the *E. coli* S10 structural gene. The DNA fragment amplified by the two *B. subtilis* oligonucleotides is about 500 bases long, and the *E. coli* S10 PCR fragment is about 420 bases long. Both PCR fragments were purified by the DNA clean-up kit from Promega.

Where appropriate, NusA from *E. coli* was added to 40 nM. The reaction components were mixed together with 500 μM (each) CTP and GTP and 5 μM UTP (no ATP) and incubated at 37°C for 10 min to allow formation of the initiation complex and incorporation of the proximal several nucleotides. In the absence of ATP, both *B. subtilis* and *E. coli* S10 promoters are able to synthesize transcripts about 4 to 11 nucleotides long. A single round of transcription elongation was then started by the addition of ATP to 500 μM, 5 μCi of [$^{32}$P]UTP, and rifampin to 10 μg/ml. Reactions were terminated at various times by the addition of 8 μl of sequencing stop mixture. The RNA products were heated at 95°C for 2 min before being loaded on an 8% sequencing gel.

**Nucleotide sequence accession number.** The S10 gene cluster sequence has been deposited in the GenBank database under accession no. U43929.

## RESULTS

**Cloning and sequencing of the *B. subtilis* S10 operon.** The sequences of the promoter distal genes of the S10 gene cluster (L29, S17, and part of L16) as well as the complete *spc* and α gene clusters of *B. subtilis* have been reported previously (1, 15, 34). To isolate DNA from the unsequenced part of the S10 gene cluster, we synthesized a primer corresponding to the known part of the L16 gene. We designed another primer based on a highly conserved amino acid motif of the L4 protein (19, 41), using codons found most frequently in other r-protein genes of *B. subtilis*. With these two primers, a 3-kb fragment was amplified by PCR from the *B. subtilis* chromosomal DNA. The fragment was sequenced and found to contain complete genes for L23, L2, S19, L22, and S3 and parts of the L4 and L16 genes (Fig. 2).

Since we wanted to determine the sequence of the DNA upstream of the S10 gene but had no information that could be used to design a primer in this region, we took an indirect PCR approach for isolating DNA from the upstream part of the S10 gene cluster. A Southern blot of *Eco*RI-digested *B. subtilis* chromosomal DNA probed with DNA from the *B. stearother-*


-35
promoter 1 -10
-35 promoter 2 -10
S10 Met Ala Lys Gln Lys Ile Arg Ile Arg Leu Lys Ala Tyr Asp His Arg Ile Leu Asp Gln Ser Ala Glu Lys Ile Val Glu
L3 Met Thr Lys Gly Ile Leu Gly Arg Lys Ile Gly
L4 Met Pro Lys Val Ala Leu Tyr Asn Gln Asn Gly Ser Thr Ala Gly Asp Ile Glu Leu Asn Ala Ser Val Phe Gly Ile
L23 Met Lys Asp Pro Arg Asp Val Leu Lys Arg Pro Val Ile Thr Glu Arg Ser Ala
L2 Met Ala Ile Lys Lys Tyr Lys Pro Ser Ser Asn Gly

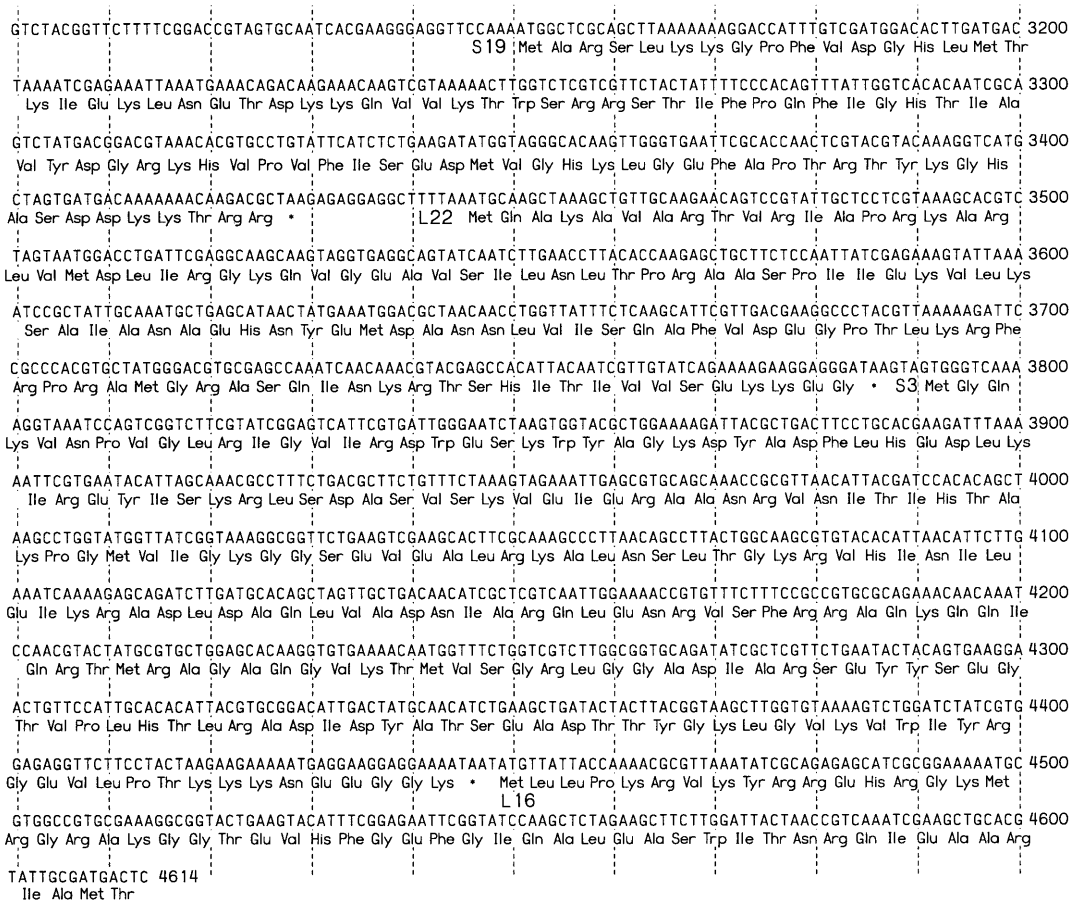


FIG. 2. Nucleotide and predicted amino acid sequences of the proximal genes of the *B. subtilis* S10 operon. The −35 and −10 regions of the two promoters upstream of the S10 gene are boxed. The fragment cloned into pPL703 for promoter activity analysis (Fig. 1A) is indicated by overlining. The region cloned in front of *lacZ* to analyze the regulatory effect of L4 (Fig. 1B) is indicated by underlining.

*mophilus* L4 gene indicated that the *B. subtilis* L4 gene is located on a 3-kb *Eco*RI fragment (data not shown). The sequence we had determined for the region downstream of the L4 gene predicted an *Eco*RI site in the L2 gene. If the organization of the S10 gene cluster in *B. subtilis* is similar to that of *E. coli*, the *Eco*RI fragment containing the L4 gene should also include the L23 and part of the L2 genes downstream as well as the promoter and the promoter-proximal S10 and L3 genes upstream. DNA was isolated from the 3-kb region of the agarose gel, purified, and ligated to a *Pst*I-*Eco*RI fragment from plasmid pLL226 (36). To PCR amplify the ligation product containing the desired *B. subtilis* DNA fragment, we used two primers, one hybridizing to the upstream flanking region of plasmid pLL226 and the other hybridizing to our previously sequenced region of the *B. subtilis* L4 gene (Fig. 2). The PCR generated a 1.5-kb fragment. Our sequence analysis showed that this fragment contains, as predicted, the genes for S10 and L3 and part of L4, as well as 300 bases upstream of the S10 gene cluster (Fig. 2).

Our studies of the *B. subtilis* S10 gene cluster confirm that the gene order is the same as that in the *E. coli* operon, although there are differences in the intergenic distances (Fig. 3). Also, *B. subtilis* S3 uses GUG as a start codon, while *E. coli* S3 initiates with AUG. The amino acid sequence identity between the homologous genes ranges from 31% for L23 to 65% for S19. L4, the regulatory protein of the *E. coli* S10 operon, has 42% identity with its *B. subtilis* homolog.

**Promoter analysis.** By visual inspection of the sequence, we identified two potential promoters which are located ~200 and ~140 bp upstream of the S10 structural gene (Fig. 2). The sequences of these two potential promoters resemble *B. subtilis* vegetative promoter sequences (24). The −10 regions of these promoters match exactly the consensus sequence (TATAAT). The upstream promoter (TTGATA) has five of six nucleotides identical to the −35 consensus (TTGACA), and the downstream promoter (TTGCAA) has four identical nucleotides.

We carried out primer extension studies to determine if these potential promoters are active in vivo. *B. subtilis* total RNA isolated from a mid-log-phase culture was used as the template. Two oligonucleotides which can hybridize to either the S10 structural gene or the untranslated leader sequence were used as primers (see Materials and Methods). With each primer we detected two major bands, in both cases corresponding in size to transcripts initiated at the two potential promoters (Fig. 4). These results imply that both promoters are functional in vivo, although the relative band intensities suggest that the downstream promoter is more active. We also analyzed the two promoters in an in vitro transcription reaction, using *E. coli* RNA polymerase. Both promoters were active in the cell-free system (data not shown; also, see below).

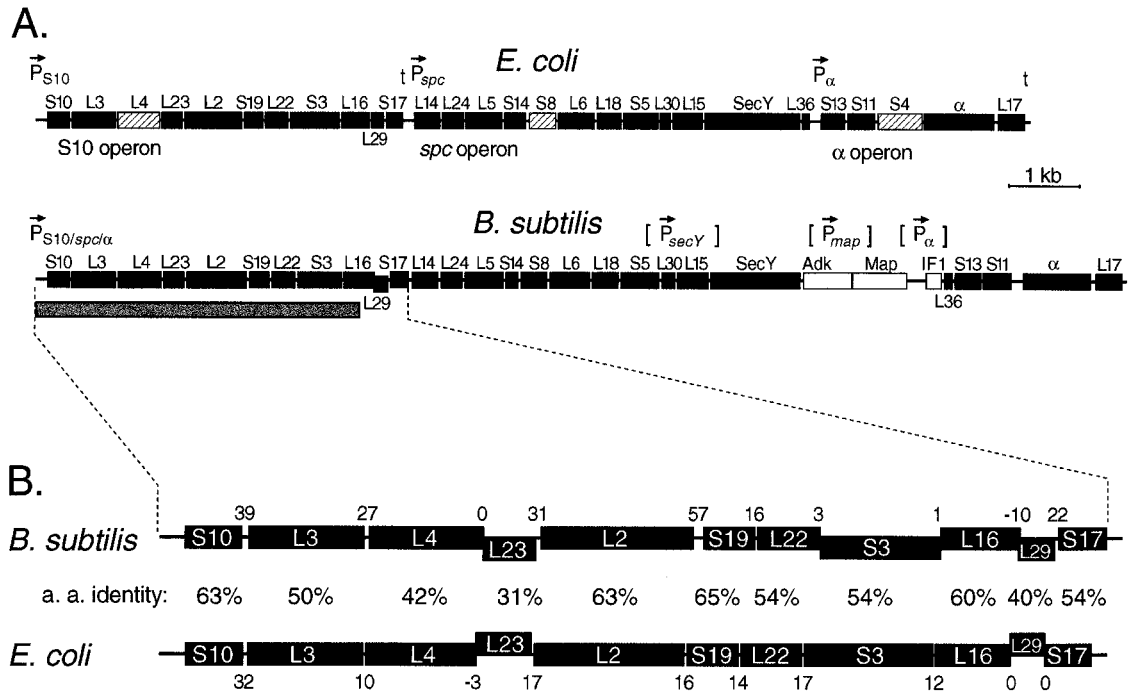


FIG. 3. (A) Comparison of the genetic organization of the S10-*spc*-α regions of *E. coli* (40) and *B. subtilis* (15, 34). The region sequenced in this study is indicated by the gray bar below the *B. subtilis* map. Promoters (P) and terminators (t) are indicated above the maps. The weak promoters identified within the *spc*-α region of *B. subtilis* (34) are shown in brackets. The genes encoding r-proteins identified as autogenous regulatory proteins in *E. coli* (40) are indicated by hatching. Genes found in the *B. subtilis* but not the *E. coli* S10-*spc*-α region are shown as white boxes. (B) Comparison of predicted amino acid sequences of the S10 operon genes of *E. coli* and *B. subtilis*. The intercistronic distances are indicated above the *B. subtilis* and below the *E. coli* maps. The amino acid sequences were aligned using the alignment program from LaserGene (DNAStar, Inc.). Identity was calculated as the percentage of amino acids in *E. coli* conserved in *B. subtilis.*

To confirm the in vivo activities of these promoters, we cloned the fragment of DNA containing these sequences (Fig. 2) into the *Bacillus* vector pPL703, which has the *B. subtilis* CAT gene without a promoter (7). The resulting plasmid, pPL703-pS10, conferred chloramphenicol resistance to *B. subtilis* SG38; transformants carrying just the vector pPL703 did not (data not shown). To better assess the strength of the *B. subtilis* S10 promoters, we determined the level of chloramphenicol resistance in cells carrying plasmid pPL703-pS10 relative to that in cells carrying plasmid pPL708, the same vector but with the *B. subtilis* SPO2 promoter (7). Cells with pPL703-pS10 could form colonies on plates with 105 μg of chloramphenicol per ml (but not 150 μg/ml), while cells carrying the reference plasmid pPL708 grew at chloramphenicol levels only up to 73 μg/ml (data not shown). We conclude that the S10 promoter sequences have a sufficiently high level of activity to be responsible for expression of the downstream r-protein genes.

**Northern analysis of mRNA from the S10, *spc*, and α clusters.** Although the S10-*spc*-α cluster of *E. coli* is organized into three transcription units, previous studies characterizing the *spc* and α clusters of *B. subtilis* concluded that expression of these genes depends on transcription initiation somewhere upstream of or within the S10 cluster (15, 34). Hence, the promoter sequences we identified upstream of the S10 gene might be responsible for transcribing the entire 30-gene S10-*spc*-α cluster. We carried out a Northern analysis of RNA purified from exponentially growing *B. subtilis* cultures, hoping to identify a 15-kb transcript containing sequences complementary to probes from all three clusters. Alternatively, if the *spc* and α cluster genes were transcribed from internal promoters, we expected to detect smaller transcripts hybridizing to a subset of the probes. However, no strong bands were detected by any of the S10, *spc*, or α cluster probes, even though a control probe complementary to the S4 gene located elsewhere on the chromosome identified a transcript of the expected size from this operon (data not shown). Thus, the Northern analysis was inconclusive. However, given the length (>15 kb) and predicted synthesis time (approximately 5 min [2]) of the putative S10-*spc*-α RNA transcript, it would not be surprising if this primary transcript was not detectable because it was processed or partially degraded before or immediately after its synthesis was completed.

Interestingly, an ~4.5-kb band could be detected by the L17 gene probe (α cluster) when RNA was prepared from stationary-phase cells. This transcript was barely detectable in mid-log-phase cells. The size of this band is approximately the same as that of a transcript initiated at the *map* promoter shown by Suh et al. (34) (Fig. 3) to have very weak activity during log phase. Our data suggest that the *map* promoter is more active during stationary phase.

**Effect of *E. coli* L4 on *B. subtilis* S10 leader-S10–*lacZ* fusion.** For unknown reasons, we were not able to construct a plasmid carrying the intact *B. subtilis* L4 gene, using either *E. coli* or *B. subtilis* as the host, although the intact *B. stearothermophilus* L4 gene has been cloned (16) and expressed in *E. coli* (41). Therefore, we could not analyze the effect of *B. subtilis* L4 on regulation of the S10 cluster in *E. coli* or *B. subtilis.*

In *E. coli* the target for L4-mediated autogenous control of the S10 operon lies within the untranslated leader and the beginning of the S10 structural gene. Leaders from closely related enterobacteria, which have predicted secondary structures that are very similar to *E. coli*'s, can be substituted for the *E. coli* RNA, showing in vivo regulation in *E. coli* by *E. coli* L4

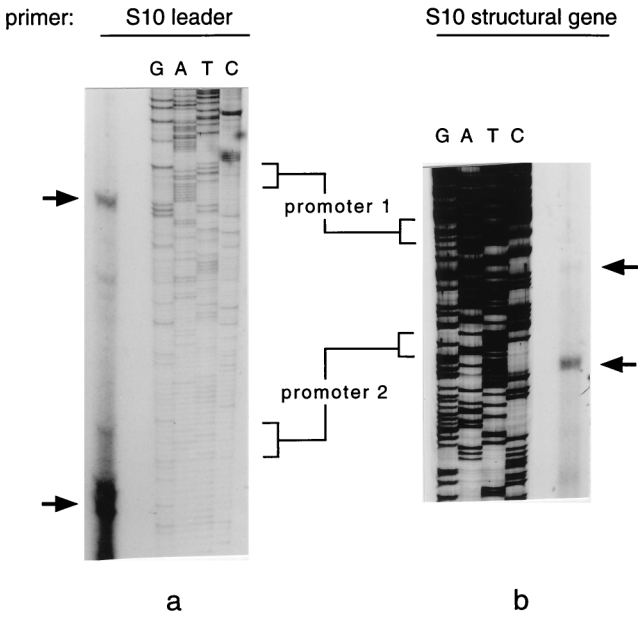


FIG. 4. Primer extension analysis of *B. subtilis* RNA. Total RNA purified from a mid-log-phase culture was used as the template. A $^{32}$P-labeled oligonucleotide hybridizing to the S10 leader (a) or S10 structural gene (b) was used as the primer for the extension reaction. The products were analyzed on an 8% gel together with a sequencing ladder. The major products are indicated by arrows. The −10 regions of the two putative promoters are indicated by brackets.

(39; also unpublished data). Conversely, L4 proteins from these same enterobacteria are also capable of regulating the *E. coli* S10 operon (41). Although the predicted S10 leader in *B. subtilis* has a length similar to that of *E. coli*'s S10 leader, there is no apparent primary sequence homology. We determined possible secondary structures of the *B. subtilis* leader using the suboptimal MFOLD computer program (35, 42). The computer-predicted structures are very different from *E. coli*'s, although one of the structures predicted to be most stable contains a large hairpin with a string of U residues that is at least superficially similar to the attenuator hairpin of *E. coli*. This hairpin structure could be formed in transcripts originating at either of the two *Bacillus* promoters (Fig. 5).

We tested the ability of the *B. subtilis* leader to function in vivo in *E. coli* as a target for *E. coli* L4 control. Since constitutive expression of "foreign" S10 leaders appears to be toxic to *E. coli* (reference 32 and our unpublished results), we cloned the leader and first 56 codons of the S10 gene from *B. subtilis* between the IPTG-inducible *trc* promoter and the *lacZ* gene. On another plasmid, *E. coli* L4 was placed under control of an arabinose-inducible promoter (Fig. 1). The constructs were then transformed together into *E. coli* (they carry compatible replication origins). The effect of L4 on expression of the S10-*lacZ* fusion genes on the *B. subtilis* (and, as a control, *E. coli*) leader-S10 construct was then monitored by pulse-labeling the cells with [$^{35}$S]methionine before and after induction with IPTG and arabinose. As expected, the expression of the S10-*lacZ* gene from both plasmids was dependent on induction with IPTG (Fig. 6, lanes 2 and 5). However, while the *E. coli* S10 leader fusion was turned off after the induction of L4 by the addition of arabinose (Fig. 6, lane 6), the *B. subtilis* construct was unaffected (Fig. 6, lane 3). These results indicate that *E. coli* L4 cannot regulate the *B. subtilis* S10 operon fusion. We have carried out similar experiments using *B. stearothermophilus* L4; this protein, which has 78% identity with *B. subtilis* L4, also fails to regulate the expression of the *B. subtilis* S10 operon in *E. coli* (data not shown).

We also analyzed the in vitro transcription of the *B. subtilis* leader, using *E. coli* RNA polymerase and r-protein L4. With an *E. coli* leader template, RNA polymerase pauses at the site of in vivo termination; this pause requires the transcription factor NusA and is stabilized by the addition of L4 (31, 38). The half-life of the pause is sensitive to the UTP concentration such that, at a low (10 μM) UTP concentration, in the presence of NusA, essentially all RNA polymerases stop at the *E. coli* S10 attenuator (Fig. 7). However, the analogous *B. subtilis* template does not respond to either NusA or L4 from *E. coli*. At 10 μM UTP, even in the presence of NusA, there is no detectable pause within the *B. subtilis* S10 leader (Fig. 7). Addition of L4 also had no effect (data not shown). Therefore, even though the RNA folding program can generate a leader structure containing a hairpin followed by a cluster of U's that is similar to the *E. coli* attenuator structure (Fig. 5), there is no evidence from the in vitro experiments that this structure, if indeed it forms, contains the determinants for NusA-stimulated pausing.

## DISCUSSION

We have sequenced a 4.5-kb region of the *B. subtilis* chromosome which contains the previously unsequenced part of the S10 cluster, encoding ribosomal proteins S10, L3, L4, L23, L2, S19, L22, and S3 and the 5′ half of L16. The *B. subtilis* S10 operon genes are organized in the same order as their *E. coli* counterparts. Alignment of predicted protein sequences of S10 operon genes for *B. subtilis* and *E. coli* shows a range of 31 to 65% amino acid identity (Fig. 3). L4, the regulatory protein of the *E. coli* S10 operon, has 42% identity with *B. subtilis* L4.

Even though the gene organization of this region is the same in *B. subtilis* and *E. coli*, the transcriptional organization is clearly different. In *E. coli*, transcription of the S10 operon is initiated from a single promoter (26, 33) and is efficiently terminated after the S17 gene (21). The *spc* operon is transcribed by its own promoter (21, 28), and the α operon is transcribed by both its own and the *spc* promoter (21, 27). In *B. subtilis* we identified two adjacent promoters upstream of the S10 operon, but there is no apparent terminator after the S17 gene (15), suggesting that in *B. subtilis*, transcription initiated at the S10 promoters extends to the end of the α operon. Furthermore, except for the internal *secY* and *map* promoters, which have very little activity during logarithmic growth (34), neither the *spc* nor the α cluster of *B. subtilis* contains recognizable promoters, implying that the S10 promoters are likely to be the major source for transcription of the *spc* and α operons.

If indeed the *B. subtilis* S10, *spc*, and α clusters consist of a single transcription unit, the predicted length of the primary transcript is over 15 kb. Such an RNA molecule is likely to be unstable, and its 5′ end may be degraded before the 3′ end is transcribed. This could explain why we failed to detect any single strong band of this size in our Northern analysis. Our ability to detect the 0.75-kb S4 transcript (10) and a 4.5-kb RNA corresponding to molecules presumably initiated at the *map* promoter (34) suggests that if there were active promoters in the *spc* and α clusters, we should have been able to detect their transcription products as well. Therefore, based on the previous studies by the Price and Henkin laboratories (15, 34) and our promoter analysis and Northern blot results, we conclude that the *B. subtilis* S10 promoters identified in this study are likely to transcribe the entire S10, *spc*, and α region.

Several lines of evidence indicate that the mechanisms for

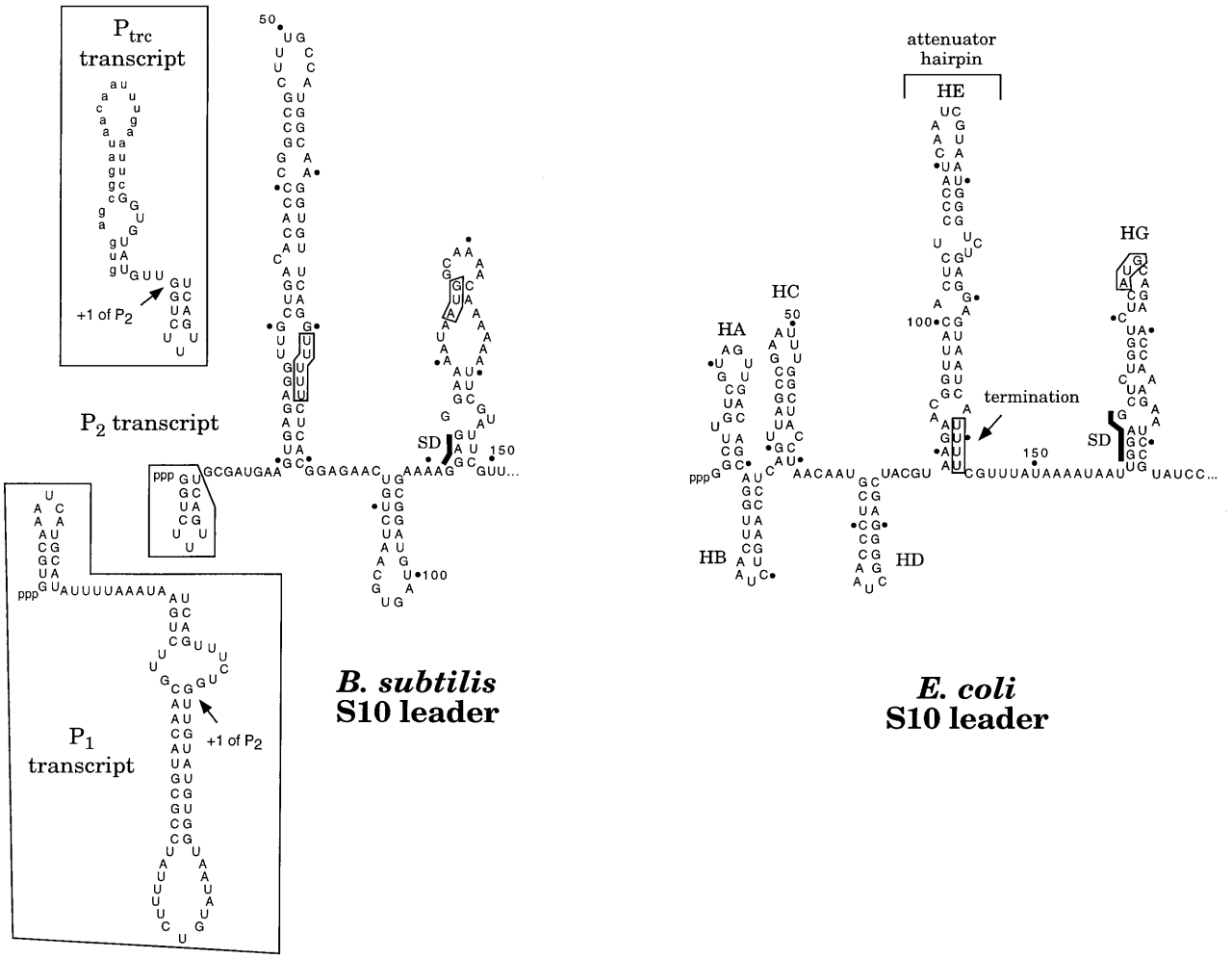


FIG. 5. Comparison of the S10 leader structures of *E. coli* and *B. subtilis*. The structure of the *E. coli* S10 leader has been described previously (33). Hairpins are designated HA through HE and HG. The structures of the various *B. subtilis* leaders were obtained by using M. Zuker's MFOLD version 2.3 (42) on the MFOLD server (URL address: http://www.ibc.wustl.edu/~zuker/rna/form1.cgi), which uses the energy rules described by Walter et al. (35). The boxed regions of the *B. subtilis* structures show predicted differences in the structures for transcripts initiated at the indicated promoters. The lowercase letters in the $P_{trc}$ transcript indicate bases transcribed from the vector sequence, including an *Eco*RI site used to construct the clone. SD, GGAG Shine-Dalgarno sequences for the S10 structural genes. The boxed AUG sequences are the initiation codons for S10. The boxed U's in the *E. coli* sequence are at the site of L4-mediated termination (36, 37). A similar run of U's in the *B. subtilis* sequence is also boxed.

regulating the expression of the *spc* and α clusters of *B. subtilis* are very different from those of *E. coli* (10, 12, 15) (see the introduction). The studies reported here suggest that regulation of the S10 cluster of *B. subtilis* will also involve a different regulatory mechanism. L4-mediated regulation of both transcription and translation of the *E. coli* S10 operon requires determinants in the S10 leader (8, 30). Inspection of the computer-predicted secondary structures of the *B. subtilis* S10 leader suggests that the region immediately upstream of the S10 gene could form a structure that is at least superficially similar to the *E. coli* attenuator hairpin. However, our in vivo studies indicate that the *B. subtilis* leader lacks sufficient similarity to be regulated by *E. coli* L4. Furthermore, transcription control by L4 in *E. coli* correlates with a NusA-dependent, L4-stimulated pause in vitro at the site where L4 stimulates in vivo termination of transcription (30, 36, 38). Such a pause is absent during in vitro transcription of the *B. subtilis* leader, albeit with *E. coli* transcription components.

We cannot rule out the possibility that the *E. coli* system cannot reproduce a *B. subtilis* L4-mediated regulatory process.

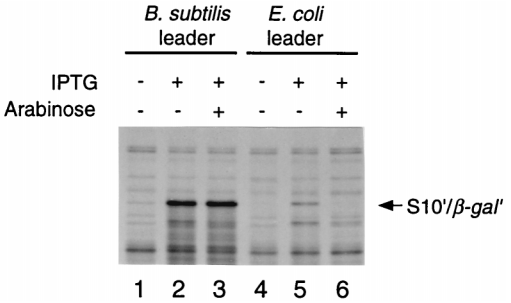


FIG. 6. Effect of L4 on expression of the *B. subtilis* S10 leader-S10–*lacZ* fusion (S10′/β-*gal*′). *E. coli* cells carrying plasmid pAra-L4 and either pACYC-Bsu or pACYC-Eco were grown in glycerol minimal media. Aliquots were pulse-labeled for 1 min with [$^{35}$S]methionine immediately before (lanes 1 and 4) or 10 min after (lanes 2 and 5) induction with IPTG. Ten minutes later arabinose was added, and after another 10 min, cells were again pulse-labeled (lanes 3 and 6).

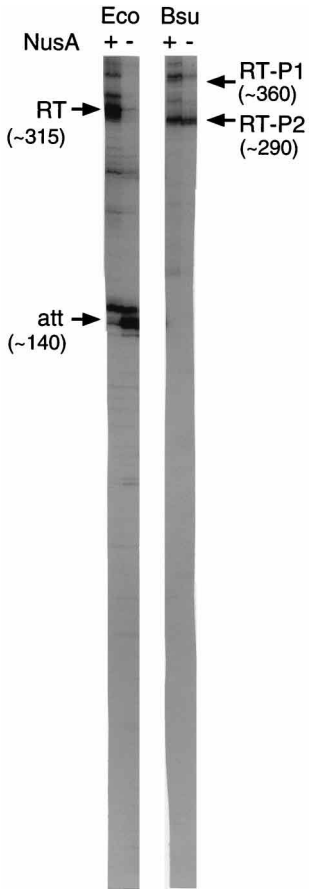


FIG. 7. Effect of *E. coli* NusA on in vitro transcription from the *B. subtilis* S10 promoters. Conditions for the transcription reactions, with PCR-amplified DNA as the template, are described in Materials and Methods. Eco, *E. coli* template; Bsu, *B. subtilis* template; RT, readthrough RNAs; att, attenuated RNAs. Both *B. subtilis* promoters (P1 and P2) were active, generating two readthrough transcripts of the expected size. The sizes of the various transcripts are given in nucleotides.

Since we were unable to clone the *B. subtilis* L4 gene, we could not directly assess whether its product regulates all or any of the r-protein genes in the S10-*spc*-α cluster of *B. subtilis*. We were able to clone the L29 and S17 genes of *B. subtilis* and found that neither protein affected expression of the *B. subtilis* S10 leader fusion in *E. coli* (data not shown). In conclusion, even though we now have identified the promoter(s) likely to drive transcription of the major *B. subtilis* r-protein gene region, the mechanisms governing the expression of the *B. subtilis* S10, *spc*, and α cluster genes are still unknown.

## ACKNOWLEDGMENTS

We thank J. McCormick, R. Losick, and P. Lovett for providing plasmids. We also thank P. Lovett, E. Rogers, and H. Schreier for technical advice.

This study was supported by grants AI-15286 and GM-54876 from the National Institutes of Health.

## REFERENCES

1. **Boylan, S. A., J.-W. Suh, S. M. Thomas, and C. W. Price.** 1989. Gene encoding the alpha core subunit of *Bacillus subtilis* RNA polymerase is cotranscribed with the genes for initiation factor 1 and ribosomal proteins B, S13, S11, and L17. J. Bacteriol. **171:**2553–2562.
2. **Bremer, H., and D. Yuan.** 1968. RNA chain growth-rate in *Escherichia coli*. J. Mol. Biol. **38:**163–180.
3. **Casadaban, M. J., and S. N. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138:**179–207.
4. **Cerretti, D. P., L. C. Mattheakis, K. R. Kearney, L. Vu, and M. Nomura.** 1988. Translational regulation of the *spc* operon in *Escherichia coli*: identification and structural analysis of the target site for S8 repressor protein. J. Mol. Biol. **204:**309–329.
5. **Cheo, D. L., K. W. Bayles, and R. E. Yasbin.** 1993. Elucidation of regulatory elements that control damage induction and competence induction of the *Bacillus subtilis* SOS system. J. Bacteriol. **175:**5907–5915.
6. **Clark, D. J., and O. Maaløe.** 1967. DNA replication and the division cycle in *Escherichia coli*. J. Mol. Biol. **23:**99–112.
7. **Duvall, E. J., D. M. Williams, P. S. Lovett, C. Rudolph, N. Vasantha, and M. Guyer.** 1981. Chloramphenicol-inducible gene expression in *Bacillus subtilis*. Gene **24:**171–177.
8. **Freedman, L. P., J. M. Zengel, R. H. Archer, and L. Lindahl.** 1987. Autogenous control of the S10 ribosomal protein operon of *Escherichia coli*: genetic dissection of transcriptional and post-transcriptional regulation. Proc. Natl. Acad. Sci. USA **84:**6516–6520.
9. **Gregory, R. R., P. B. F. Cahill, D. L. Thurlow, and R. A. Zimmermann.** 1988. Interaction of *Escherichia coli* ribosomal protein S8 with its binding sites in ribosomal RNA and messenger RNA. J. Mol. Biol. **204:**295–307.
10. **Grundy, F. J., and T. M. Henkin.** 1990. Cloning and analysis of the *Bacillus subtilis rpsD* gene, encoding ribosomal protein S4. J. Bacteriol. **172:**6372–6379.
11. **Grundy, F. J., and T. M. Henkin.** 1991. The *rpsD* gene, encoding ribosomal protein S4, is autogenously regulated in *Bacillus subtilis*. J. Bacteriol. **173:**4595–4602.
12. **Grundy, F. J., and T. M. Henkin.** 1992. Characterization of the *Bacillus subtilis rpsD* regulatory target site. J. Bacteriol. **174:**6763–6770.
13. **Henkin, T. M.** 1993. Ribosomal structure and genetics, p. 669–682. *In* A. L. Sonenshein, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis* and other gram-positive bacteria: biochemistry, physiology, and molecular genetics. American Society for Microbiology, Washington, D.C.
14. **Henkin, T. M., and G. H. Chambliss.** 1984. Genetic mapping of a mutation causing an alteration in *Bacillus subtilis* ribosomal protein S4. Mol. Gen. Genet. **193:**364–369.
15. **Henkin, T. M., S. H. Moon, L. C. Mattheakis, and M. Nomura.** 1989. Cloning and analysis of the *spc* ribosomal protein operon of *Bacillus subtilis*: comparison with the spc operon of *Escherichia coli*. Nucleic Acids Res. **17:**7469–7486.
16. **Herwig, S., V. Kruft, and B. Wittmann-Liebold.** 1992. Primary structures of ribosomal proteins L3 and L4 from *Bacillus stearothermophilus*. Eur. J. Biochem. **207:**877–885.
17. **Jinks-Robertson, S., and M. Nomura.** 1982. Ribosomal protein S4 acts in *trans* as a translational repressor to regulate expression of the α operon in *Escherichia coli*. J. Bacteriol. **151:**193–202.
18. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
19. **Li, X., L. Lindahl, and J. M. Zengel.** 1996. Ribosomal protein L4 from *Escherichia coli* utilizes different determinants for its structural and regulatory functions. RNA **2:**24–37.
20. **Lindahl, L., R. Archer, and J. M. Zengel.** 1983. Transcription of the S10 ribosomal protein operon is regulated by an attenuator in the leader. Cell **33:**241–248.
21. **Lindahl, L., F. Sor, R. H. Archer, M. Nomura, and J. M. Zengel.** 1990. Transcriptional organization of the S10, spc and α operons of *Escherichia coli*. Biochim. Biophys. Acta **1050:**337–342.
22. **Lindahl, L., and J. M. Zengel.** 1979. Operon-specific regulation of ribosomal protein synthesis in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **76:**6542–6546.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Moran, C. P., Jr., N. Lang, S. F. J. Legrice, G. Lee, M. Stephens, A. L. Sonenshein, J. Pero, and R. Losick.** 1982. Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. **186:**339–346.
25. **Nomura, M.** 1988. Feedback regulation of ribosome biosynthesis in *Escherichia coli*, p. 97–107. *In* D. Bissell, G. Deho, G. Sironi, and A. Torriani (ed.), Gene expression and regulation: the legacy of Luigi Gorini. Elsevier, Amsterdam, The Netherlands.
26. **Olins, P. O., and M. Nomura.** 1981. Regulation of the S10 ribosomal protein operon in *E. coli*: nucleotide sequence at the start of the operon. Cell **26:**205–211.
27. **Post, L. E., A. E. Arfsten, G. R. Davis, and M. Nomura.** 1980. DNA sequence of the promoter region for the α ribosomal protein operon in *Escherichia coli*. J. Biol. Chem. **255:**4653–4659.
28. **Post, L. E., A. E. Arfsten, F. Reusser, and M. Nomura.** 1978. DNA sequences of promoter regions for the *str* and *spc* ribosomal protein operons in *Escherichia coli*. Cell **15:**215–229.
29. **Sanger, F., S. Nicklen, and A. R. Coulsen.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
30. **Sha, Y., L. Lindahl, and J. M. Zengel.** 1995. RNA determinants required for

L4-mediated attenuation control of the S10 r-protein operon of *Escherichia coli*. J. Mol. Biol. **245:**486–498.
31. **Sha, Y., L. Lindahl, and J. M. Zengel.** 1995. Role of NusA in L4-mediated attenuation control of the S10 r-protein operon of *Escherichia coli*. J. Mol. Biol. **245:**474–485.
32. **Shen, P.** 1991. RNA secondary structure and the regulation of the S10 ribosomal protein operon in *Escherichia coli*. Ph.D. thesis. University of Rochester, Rochester, N.Y.
33. **Shen, P., J. M. Zengel, and L. Lindahl.** 1988. Secondary structure of the leader transcript from the *Escherichia coli* S10 ribosomal protein operon. Nucleic Acids Res. **16:**8905–8924.
34. **Suh, J. W., S. A. Boylan, S. H. Oh, and C. W. Price.** 1996. Genetic and transcriptional organization of the *Bacillus subtilis spc*-α region. Gene **169:** 17–23.
35. **Walter, A. E., D. H. Turner, J. Kim, M. H. Lyttle, P. Müller, D. H. Mathews, and M. Zuker.** 1994. Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. Proc. Natl. Acad. Sci. USA **91:**9218–9222.
36. **Zengel, J. M., and L. Lindahl.** 1990. *Escherichia coli* ribosomal protein L4 stimulates transcription termination at a specific site in the leader of the S10 operon independent of L4-mediated inhibition of translation. J. Mol. Biol. **213:**67–78.
37. **Zengel, J. M., and L. Lindahl.** 1990. Ribosomal protein L4 stimulates in vitro termination of transcription at a NusA-dependent terminator in the S10 operon leader. Proc. Natl. Acad. Sci. USA **87:**2675–2679.
38. **Zengel, J. M., and L. Lindahl.** 1992. Ribosomal protein L4 and transcription factor NusA have separable roles in mediating termination of transcription within the leader of the S10 operon of *E. coli*. Genes Dev. **6:**2655–2662.
39. **Zengel, J. M., and L. Lindahl.** 1993. Regulation of the eleven gene S10 ribosomal protein operon by the 50S subunit protein L4, p. 131–138. *In* K. Nierhaus, F. Franceschi, and A. P. Subramanian (ed.), The translational apparatus: structure, function, regulation, evolution. Plenum Press, New York, N.Y.
40. **Zengel, J. M., and L. Lindahl.** 1994. Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli*. Prog. Nucleic Acid Res. Mol. Biol. **47:**331–370.
41. **Zengel, J. M., D. Vorozheikina, X. Li, and L. Lindahl.** 1995. Regulation of the *Escherichia coli* S10 ribosomal protein operon by heterologous L4 ribosomal proteins. Biochem. Cell Biol. **73:**1105–1112.
42. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244:**48–52.