Genetic Regulation of the yefM-yoeB Toxin-Antitoxin Locus of Streptococcus pneumoniae

Wai Ting Chan,¹,² Concha Nieto,² Jennifer Ann Harikrishna,¹ Seok Kooi Khoo,³ Rofina Yasmin Othman,¹ Manuel Espinosa,²,* and Chew Chieng Yeo⁴,*

Centre for Research in Biotechnology for Agriculture (CEBAR) and Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia; Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Ramiro de Maeztu 9, E28040-Madrid, Spain; Institute of Molecular Biology, Academia Sinica, Taipei 115, Taiwan; and Faculty of Agriculture and Biotechnology, Universiti Sultan Zainal Abidin, Jalan Sultan Mahmud, 20400 Kuala Terengganu, Malaysia

Received 28 April 2011/Accepted 5 July 2011

Type II (proteic) toxin-antitoxin systems (TAS) are ubiquitous among bacteria. In the chromosome of the pathogenic bacterium Streptococcus pneumoniae, there are at least eight putative TAS, one of them being the yefM-yoeBSpn operon studied here. Through footprinting analyses, we showed that purified YefMSpn antitoxin and the YefM-YoeBSpn TA protein complex bind to a palindrome sequence encompassing the −35 region of the main promoter (P<sup>yefM2</sup>) of the operon. Thus, the locus appeared to be negatively autoregulated with respect to P<sup>yefM1</sup>, since YefMSpn behaved as a weak repressor with YoeBSpn as a corepressor. Interestingly, a BOX element, composed of a single copy (each) of the boxA and boxC subelements, was found upstream of promoter P<sup>yefM2</sup>. BOX sequences are pneumococcal, perhaps mobile, genetic elements that have been associated with bacterial processes such as phase variation, virulence regulation, and genetic competence. In the yefM-yoeBSpn locus, the boxAC element provided an additional weak promoter, P<sup>boxAC</sup>, upstream of P<sup>yefM2</sup> which was not regulated by the TA proteins. In addition, transcriptional fusions with a lacZ reporter gene showed that P<sup>yefM1</sup> was constitutive albeit weaker than P<sup>yefM2</sup>. Intriguingly, the coupling of the boxAC element to P<sup>yefM1</sup> and yefM<sub>Spn</sub> in cis (but not in trans) led to transcriptional activation, indicating that the regulation of the yefM-yoeBSpn locus differs somewhat from that of other TA loci and may involve as yet unidentified elements. Conservation of the boxAC sequences in all available sequenced genomes of S. pneumoniae which contained the yefM-yoeBSpn locus suggested that its presence may provide a selective advantage to the bacterium.

Detection of gene variations and regulatory circuits in microbial systems is critical for our knowledge of the evolvability of bacterial species, with one of the driving forces underlying bacterial evolution being stressful environmental conditions (2). Important operative pieces intervening in bacteria that cope with stress are the so-called toxin-antitoxin systems (TAS). Among them, the most studied are those belonging to the type II (proteic) TAS. These usually comprise two cotranscribed genes that encode an unstable antitoxin and a stable toxin. The antitoxin binds to and neutralizes its cognate toxin through protein-protein interactions. Under environmental stress conditions which cause the induction of endogenous proteases, the balance between toxin and antitoxin is shifted and the toxin is released from the TA complex due to protease-mediated degradation of the antitoxin. Thus, the unbound toxin is now free to exert its effect of poisoning the cell machinery so that growth is arrested (8, 10, 54).

The discovery of the mazEF (or chpAB) chromosomally encoded TAS in Escherichia coli (28) led to the proposal that, when encoded in chromosomes, these loci could function as mediators of programmed cell death (1, 9). Unfavorable growth conditions could trigger this pathway, and as a consequence, a subpopulation of bacterial cells would die. Death of these cells would do the following: (i) provide food for the remaining population (altruistic behavior), (ii) serve as a defense mechanism to restrict phage spreading (protection against incoming DNA), and (iii) act as a mechanism to eliminate cells with deleterious mutations (preservation of the gene pool). It would seem that at least in the case of mazEF, cell death is a population-dependent phenomenon requiring a quorum-sensing molecule, termed extracellular death factor, which is a linear pentapeptide (NNWNN) important for mazEF-mediated killing activity (18). Another line of independent investigations led to the different view that TAS would function as modulators of the global levels of translation during environmental stress and that the toxin-mediated inhibition of protein synthesis led to reversible cell stasis rather than cell death (4, 38). In the case of the pneumococcal RelBE<sub>Spn</sub> TAS, it was shown that growth cessation could be made irreversible if exposure to the toxin was prolonged for periods longer than 4 h, so that the toxin-exposed cells could not be rescued by triggering antitoxin synthesis (35–36). Other hypotheses to explain the ubiquity and redundancy of the bacterial TAS (37) include, in addition to the above, a role of TAS in persistence, a process that enables a fraction of a bacterial population to survive prolonged exposure to antibiotics (i.e.,
antibiotic tolerance) by entering a state of dormancy against DNA loss (53). Perhaps the functions of TAS cannot be generalized since they may have multiple biological roles that are dependent on the nature of the toxin, their location on the genome, and other yet-to-be-discovered factors.

Concerning regulation of the TAS, it has been shown that in most cases expression of TAS seems to be controlled by self-regulatory transcriptional mechanisms in which the antitoxin would act as a repressor by binding to an operator site that overlaps the promoter. In most instances, the cognate toxin acts as a very efficient corepressor of transcription (10–11, 26, 34).[H9275]

our laboratories have been studying some of the TAS encoded by the chromosome of the Gram-positive bacterium Streptococcus pneumoniae (the pneumococcus), an important human pathogen responsible for over 2,000,000 human deaths per year (25). Sequence analyses performed on the available pneumococcal genomes have shown that the bacterium harbors at least eight TAS homologues: relBE1spn, relBE2spn, yefM-yoeBSpn, higAB, phd-doc, pezAT, tasAB, and hicAB (36). Of these, three TAS, namely, relBE2spn, yefM-yoeBSpn, and pezAT, have been studied and characterized as bona fide TAS, whereas relBE1spn was shown to be not functional (16, 34–35).

This number of pneumococcal TAS could be an underestimation, since database mining of sequenced microbial genomes indicates an abundance of TA loci, particularly in free-living microorganisms (37), and the number of reported TAS has been steadily increasing (24). A database containing most known and predicted TAS (48) is available at http://bioinfo.mml.sjtu.edu.cn/TADB/.

Here we present a study on the regulation of the pneumococcal yefM-yoeBSpn locus, which is transcribed as a single unit. Through DNA binding and footprinting assays, in conjunction with transcriptional fusions, we have demonstrated that control of expression of this operon is at three levels. First, in the case of other TAS, the YefMSpn, antitoxin can bind to DNA and repress transcription from its own promoter, P_{yefM^2}. Second, the YoeBSpn toxin exerts further repression of transcription from the same promoter both in cis and in trans by binding to the YefMSpn antitoxin. Third and most distinctly, we have detected the insertion of a pneumococcal BOX element (comprising boxA and boxC subelements) upstream of P_{yefM^2} which has led to the generation of a new and YefM-YoeBSpn-unregulated promoter, P_{yefM}. The BOX was conserved in all sequenced pneumococcal strains but absent in other yefM-yoeB homologues of different bacteria. Thus, BOX seems to have contributed to introducing a new level of complexity to the regulation of the yefM-yoeBSpn locus by increasing the overall basal transcription level. BOX elements are repeated sequences (up to 125 copies per genome) which have been associated with genetic competence, virulence (17), and phenotypic phase variation (45) in S. pneumoniae. Although little is known on their contribution to the genetics of S. pneumoniae (6), it would appear that the hyperrecombinogenic nature of S. pneumoniae (12) may provide the bacterium with selective advantages when subjected to stressful conditions.

MATERIALS AND METHODS

Bacterial strains, transformation, and growth conditions. E. coli Top10 [F- mcrBC Δmrr-hsdS-M15 ΔrecA1 ΔlacZΔM15 ΔproAB ΔproC AB △(leuB7697 galU galK rpsL) F′ lacZΔM15 galK rpsL endA1 (ori V41) supG (Strf)] (Invitrogen) was used as the host for cloning experiments. E. coli BL21-CodonPlus(DE3)-RIL [E. coli B (F- ompT hsdS (B^+ M^-) dcm +) Tol^+ galE (DE3) endA1 (HsdS) galK (DE3) H2 (dey adv1 YeaW Cam^+)] (Stratagene) was used as the host for the overexpression of genes cloned into the pET28a expression vectors. E. coli cultures were grown in Luria-Bertani medium (Difco) at 37°C. Liquid cultures were grown in an orbital shaker incubator at 250 rpm. When necessary, growth medium for E. coli was supplemented with antibiotics at the concentrations: ampicillin, 100 μg/ml; chloramphenicol, 34 μg/ml; and kanamycin, 50 μg/ml. S. pneumoniae R6 cells (21) were grown in AGCH medium (20) supplemented with 1% sucrose and 0.25% yeast extract. All S. pneumoniae cultures were grown at 37°C.

DNA manipulations, sequencing, and sequence data analysis. DNA manipulations and other molecular biology techniques were carried out using established protocols (46). Genomic DNA from S. pneumoniae was prepared using the Wizard genomic DNA purification kit (Promega). DNA fragments and PCR products were purified using the GFX purification kit from Amersham Biosciences. DNA sequencing was carried out using the BigDye Terminator cycle sequencing ready reaction kit and an ABI Prism 377 DNA sequencer (Applied Biosystems Inc.). DNA sequences were compiled and analyzed using the Lasergene sequence analysis software program (DNASTar). Comparison of nucleotide and amino acid sequence data were carried out using the Lasergene sequence analysis program and DNAStar.

RNA preparation and RT-PCR. Total RNA samples were prepared from 10 ml S. pneumoniae cultures using an RNasy Midi kit (Qiagen) according to the manufacturer's procedure. The total RNA concentrations were determined byTT quantitation of the RNA samples, and the RNA quality was analyzed on a 1% agarose gel. Total RNA (400 ng) and 10 μM gene-specific primers were mixed and heated for 5 min at 70°C and then placed on ice. This mixture was added to a 20 μl reverse transcriptase (RT) reaction mixture containing 4 μl of cDNA synthesis buffer, 5 μM dithiothreitol (DTT), 40 U RNAse OUT, 1 mM deoxyadenosine triphosphate (dATP), 15 U T8 DNA polymerase (United States Biochemical) by annealing the yefM-near or the yefM-far primer to the pMBH13 denatured plasmid DNA, respectively. A mixture of dATP, dGTP, dCTP, dTTP (100 μM each), and 10 μM [32P]dCTP was then added. The primers were then extended with ThermoScript reverse transcriptase (Invitrogen) at 50°C or at 55°C for 30 min. After that, 80 μl of Tris-EDTA (TE) was added, followed by phenol-chloroform extraction and then ethanol precipitation. The pellet was dissolved in 8 μl of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol) and loaded onto an 8% urea polyacrylamide gel (7 M) along with the dideoxy sequencing DNA ladder. The dideoxy sequencing DNA ladder was obtained using Sequenase version 2.0 DNA polymerase (United States Biochemical) by annealing the yefM-near or the yefM-far primer to the pMBH13 (34) denatured plasmid DNA, respectively.

Construction of lacZ transcriptional fusions and use of β-galactosidase assays to detect promoter activity. The promoter-probe vector pQF52 (30) was used to construct lacZ transcriptional fusions for the detection of promoter activity (Table 1). pQF52, which is a derivative of pACYC184, was used to construct recombinant plasmids for overexpression of YefMSpn and YefM-YoeBSpn under the control of the arabinose-inducible PAAD promoter. Details of the construction of all recombinant plasmids as well as the primer sequences are listed in Table 2 and in Table S2 in the supplemental material.

To express transcriptional fusions and use β-galactosidase assays to detect promoter activity, the recombinant plasmids were transformed into E. coli Top10 as described previously (39), by annealing at 65°C for 5 min. The transformed E. coli Top10 cells were then incubated in LB media containing antibiotics at the following concentrations: ampicillin, 100 μg/ml; chloramphenicol, 34 μg/ml; and kanamycin, 50 μg/ml. S. pneumoniae R6 cells (21) were grown in AGCH medium (20) supplemented with 1% sucrose and 0.25% yeast extract. All S. pneumoniae cultures were grown at 37°C.

DNA manipulations, sequencing, and sequence data analysis. DNA manipulations and other molecular biology techniques were carried out using established protocols (46). Genomic DNA from S. pneumoniae was prepared using the Wizard genomic DNA purification kit (Promega). DNA fragments and PCR products were purified using the GFX purification kit from Amersham Biosciences. DNA sequencing was carried out using the BigDye Terminator cycle sequencing ready reaction kit and an ABI Prism 377 DNA sequencer (Applied Biosystems Inc.). DNA sequences were compiled and analyzed using the Lasergene sequence analysis software program (DNASTar). Comparison of nucleotide and amino acid sequence data were carried out using the Lasergene sequence analysis program and DNAStar.

RNA preparation and RT-PCR. Total RNA samples were prepared from 10 ml S. pneumoniae cultures using an RNasy Midi kit (Qiagen) according to the manufacturer's procedure. The total RNA concentrations were determined byTT quantitation of the RNA samples, and the RNA quality was analyzed on a 1% agarose gel. Total RNA (400 ng) and 10 μM gene-specific primers were mixed and heated for 5 min at 70°C and then placed on ice. This mixture was added to a 20 μl reverse transcriptase (RT) reaction mixture containing 4 μl of cDNA synthesis buffer, 5 μM dithiothreitol (DTT), 40 U RNAse OUT, 1 mM deoxyadenosine triphosphate (dATP), 15 U T8 DNA polymerase (United States Biochemical) by annealing the yefM-near or the yefM-far primer to the pMBH13 denatured plasmid DNA, respectively. A mixture of dATP, dGTP, dCTP, dTTP (100 μM each), and 10 μM [32P]dCTP was then added. The primers were then extended with ThermoScript reverse transcriptase (Invitrogen) at 50°C or at 55°C for 30 min. After that, 80 μl of Tris-EDTA (TE) was added, followed by phenol-chloroform extraction and then ethanol precipitation. The pellet was dissolved in 8 μl of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol) and loaded onto an 8% urea polyacrylamide gel (7 M) along with the dideoxy sequencing DNA ladder. The dideoxy sequencing DNA ladder was obtained using Sequenase version 2.0 DNA polymerase (United States Biochemical) by annealing the yefM-near or the yefM-far primer to the pMBH13 (34) denatured plasmid DNA, respectively.

Construction of lacZ transcriptional fusions and use of β-galactosidase assays to detect promoter activity. The promoter-probe vector pQF52 (30) was used to construct lacZ transcriptional fusions for the detection of promoter activity (Table 1). pQF52, which is a derivative of pACYC184, was used to construct recombinant plasmids for overexpression of YefMSpn and YefM-YoeBSpn under the control of the arabinose-inducible PAAD promoter. Details of the construction of all recombinant plasmids as well as the primer sequences are listed in Table 2 and in Table S2 in the supplemental material.

Downloaded from http://jb.asm.org/ on July 22, 2018 by guest
Recombinant pQF52 and pLNBAD plasmids were transformed into E. coli Top10, and transformants were assayed for β-galactosidase activity according to the method of Miller (31) using SDS and chloroform to permeabilize the cells. The β-galactosidase assay for each construct was repeated eight times, and the mean value obtained was used for further analysis.

**Overexpression and purification of recombinant (His)_6-YefMSpn and YoeB in E. coli.** The YefMSpn antitoxin protein was obtained from pEMH10 (34), whereas the YefM-YoeBSpn TA complex was obtained from the recombinant plasmid constructed in this study (pET28a_HisYefMYoeB). To construct a recombinant plasmid with YefM-YoeBSpn, this operon was amplified using the primer pairs pETyefM-F and pETyoeB-R (Table 2; see also Table S2 in the supplemental material), and the amplified fragment was digested with NdeI and BamHI prior to ligation into the NdeI-BamHI site downstream of the isopropyl-β-D-1-thiogalactopyranoside (IPTG)-inducible T7 promoter of the pET28a plasmid (Novagen). Both recombinant plasmids were transformed separately into E. coli BL21-CodonPlus(DE3)-RII (Stratagene). To obtain adequate amounts of protein for this study, 8 ml of an overnight culture of E. coli BL21-CodonPlus(DE3)-RII harboring the pEMH10 or pET28a_HisYefMYoeB recombinants (Table 2), was diluted into 800 ml of fresh Luria-Bertani broth supplemented with 50 μg/ml kanamycin and 34 μg/ml chloramphenicol and allowed to grow until reaching an optical density at 600 nm (OD600) of ∼0.4 with shaking at 250 rpm and at 37°C. The cell cultures were cooled to 30°C before IPTG was added (to a final concentration of 200 μM) 15 min after addition of IPTG. After 2 h, the cells were harvested at 6,500 × g at 4°C for 20 min. The cell pellets were then resuspended in 40 ml of 1× His buffer (10 mM Tris, pH 7.6, 1 M NaCl, 5 mM β-mercaptoethanol, and 5% glycerol) containing 10 mM imidazole and 10 μM protease inhibitor mix (General Electric Healthcare). All purification steps were performed at 4°C. The cell suspension was then subjected to a French press (Constant System) twice, and the supernatant was separated from the cell debris and unbroken cells by centrifugation at 17,000 × g for 20 min. The crude lysate was then loaded into a chromatography column packed with HIS-Select nickel affinity gel (Sigma) with a flow rate of 45 ml/h. HIS-Select nickel affinity gel is an immobilized metal-ion affinity chromatography product which is a quadridentate chelate on 6% beaded agarose charged with nickel and is selective for recombinant proteins with histidine tags and exhibits low nonspecific binding of other proteins. The column was then extensively washed with 100 ml of 1× His buffer A (10 mM Tris, pH 7.6, 0.3 M NaCl, 5 mM β-mercaptoethanol, and 5% glycerol) containing 10 mM imidazole with a flow rate of 45 ml/h. The histidine-tagged YefMSpn and YoeBSpn proteins were eluted from the column using 1× His buffer with 250 mM imidazole at a flow rate of 45 ml/h. The protein fractions were collected and analyzed using 16% SDS-PAGE. One band was detected on July 22, 2018 by guest

| Construct | pHefM | pHefM2 | In trans yefMSpn yoeBSpn | yefMSpn | yoeBSpn | Ratio \(a\) |
|-----------|------|-------|-----------------------|--------|--------|----------|
| pQF52     | No   | No    | No                    | No     | No     | 0.00     |
| pQF_P2    | No   | Yes   | No                    | No     | No     | 1.00     |
| pQF_P1    | Yes  | No    | No                    | No     | No     | 0.07     |
| pQF_P1P2  | Yes  | Yes   | Yes                   | No     | No     | 0.57     |
| pQF_nP    | No   | No    | No                    | No     | No     | 0.00     |
| pQF_P2 and pLN_yM | No | Yes | Yes | No | No | 0.93 |
| pQF_P2 and pLN_yMyB | No | Yes | Yes | Yes | No | 0.13 |
| pQF_P2yM  | No   | Yes   | Yes                   | Yes    | No     | 0.40     |
| pQF_P2yMyB | No | Yes | Yes | No | Yes | 0.13 |
| pQF_P1P2 and pLN_yM | Yes | Yes | Yes | No | No | 0.50 |
| pQF_P1P2 and pLN_yMyB | Yes | Yes | Yes | Yes | No | 0.48 |
| pQF_P1P2yM | Yes | Yes | Yes | No | Yes | 1.36 |
| pQF_P1P2yMyB | Yes | Yes | Yes | Yes | Yes | 1.13 |
| pQF_M1yM  | Yes  | Yes   | Yes                   | No     | Yes    | 1.89     |
| pQF_M1yMyB | Yes | Yes | Yes | No | Yes | 2.03 |
| pQF_M2yM  | Yes  | Yes   | No                    | No     | Yes    | 1.87     |
| pQF_M2yMyB | Yes | Yes | Yes | No | Yes | 2.07 |
| pQF_P1 and pLN_yM | Yes | No | Yes | No | Yes | 0.07 |
| pQF_P1 and pLN_yMyB | Yes | No | Yes | Yes | Yes | 0.07 |
| pQF_P1yM  | Yes  | No    | Yes                   | No     | Yes    | 0.07     |
| pQF_P1yMyB | Yes | No | No | Yes | Yes | 1.23 |
| pQF_nP and pLN_yM | No | No | No | Yes | Yes | 0.00 |
| pQF_nP and pLN_yMyB | No | No | No | Yes | Yes | 0.00 |
| pQF_nPyM  | No   | No    | No                    | No     | Yes    | 1.08     |
| pQF_nPyMyB | No | No | No | Yes | Yes | 0.00 |
| pQF_M1yM  | Yes  | No    | Yes                   | No     | Yes    | 0.98     |
| pQF_M1yMyB | Yes | No | Yes | Yes | Yes | 1.28 |
| pQF_M2yB  | Yes  | No    | Yes                   | No     | Yes    | 0.98     |
| pQF_M2yMyB | Yes | No | Yes | Yes | Yes | 1.28 |

\(a\) The ratios were calculated by normalization of the β-galactosidase activity levels with that of pQF_P2, which is 122 ± 36 MU.
### TABLE 2. Recombinant constructs in this work

| Recombinant construct | E. coli host strain | Primer pair(s) | Description |
|-----------------------|---------------------|----------------|-------------|
| pQF_P1P2             | TOP10               | P1-F and P2-R  | 237 bp upstream of yefM cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene |
| pQF_P1P2yM           | TOP10               | P1-F and yefM-R| yefM cloned along with 237 bp of its upstream sequence cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene |
| pQF_P1P2yMyB         | TOP10               | P1-F and yoeB-R| yefM-yoeB cloned along with 237 bp of its upstream sequence cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene |
| pQF_P1              | TOP10               | P1-F and mP1-R; mP2-F and P2-R | 237 bp upstream of yefM cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene; the −10 region of P<sub>yefM</sub> and the nucleotides flanking this region were substituted from 5'-tgTATAAAT-3' to 5'-ctgcAg-3' |
| pQF_P1yM            | TOP10               | P1-F and mP1-R; mP2-F and yefM-R | yefM cloned along with 237 bp of its upstream sequence cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene; the −10 region of P<sub>yefM</sub> and the nucleotides flanking this region were substituted from 5'-tgTATAAAT-3' to 5'-ctgcAg-3' |
| pQF_P1yMyB          | TOP10               | P1-F and mP1-R; mP2-F and yoeB-R | yefM-yoeB cloned along with 237 bp of its upstream sequence cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene; the −10 region of P<sub>yefM</sub> and the nucleotides flanking this region were substituted from 5'-tgTATAAAT-3' to 5'-ctgcAg-3' |
| pQF_P2              | TOP10               | P2-F and P2-R  | 87 bp upstream of yefM cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene |
| pQF_P2yM            | TOP10               | P2-F and yefM-R| yefM cloned along with 87 bp of its upstream sequence cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene |
| pQF_P2yMyB          | TOP10               | P2-F and yoeB-R| yefM-yoeB cloned along with 87 bp of its upstream sequence cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene |
| pQF_nP              | TOP10               | P2-F and mP1-R; mP2-F and P2-R | 87 bp upstream of yefM cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene; the −10 region of P<sub>yefM</sub> and the nucleotides flanking this region were substituted from 5'-tgTATAAAT-3' to 5'-ctgcAg-3' |
| pQF_nPyM            | TOP10               | P2-F and mP1-R; mP2-F and yefM-R | yefM cloned along with 87 bp of its upstream sequence cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene; the −10 region of P<sub>yefM</sub> and the nucleotides flanking this region were substituted from 5'-tgTATAAAT-3' to 5'-ctgcAg-3' |
| pQF_nPyMyB          | TOP10               | P2-F and mP1-R; mP2-F and yoeB-R | yefM-yoeB cloned along with 87 bp of its upstream sequence cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene; the −10 region of P<sub>yefM</sub> and the nucleotides flanking this region were substituted from 5'-tgTATAAAT-3' to 5'-ctgcAg-3' |
| pQF_yM              | TOP10               | yefM-F and yefM-R| yefM cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene |
| pQF_CyM             | TOP10               | CyeM-F and yefM-R| 195 bp upstream of yoeB cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene |
| pQF_yB              | TOP10               | yoeB-F and yoeB-R| yoeB cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene |
| pQF_CyB             | TOP10               | CyeB-F and yoeB-R| 159 bp upstream of yoeB cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene |
| pQF_M1yM            | TOP10               | yefM7-F and yefM7-R| yefM cloned along with 237 bp of its upstream sequence cloned into NcoI-BamHI site of pQF52, preceding the promoterless lacZ gene; using site-directed mutagenesis, two base pairs were substituted, which led to the 7th codon for YefMSpn, "S," being replaced by an amber stop codon |

Continued on following page
Edman degradation method at the Protein Chemistry Facility of the Centro de
Investigaciones Biológicas.

EMSA, DNase I, and hydroxyl radical footprinting assays. DNA fragments
containing the entire upstream sequence were PCR amplified using each of the
[yr]-ATP-labeled oligonucleotides and nonlabeled oligonucleotides: labeled
PS1-F and non-labeled PS2-R (see Table S2 in the supplemental material), as
well as nonlabeled PS1-F and labeled PS2-R. EMSA was carried out in the
following mixture to a final volume of 5
ml: 1× binding buffer (100 mM Tris, pH
7.6, 5 mM EDTA, 5 mM DTT, and 5% glycerol), 3,000 cpm of labeled DNA, 10
ng/ml of heparin, and increasing concentrations of either YefMSpn or YefM-
YoeBSpn. The reaction mix was incubated for 20 min at room temperature before
being loaded onto a 5% polyacrylamide gel and run with 0.5× Tris-borate-
EDTA (TBE) buffer. The gel was then transferred to a film cassette and exposed
for 70°C until the desired image intensity was obtained.

The hydroxyl radical footprinting reagent was prepared by mixing equal vol-
umes of 0.6% H2O2, 20 mM sodium ascorbate, and FeCl2–EDTA (equal volumes
of 0.4 mM FeCl2– in sterile distilled water and 0.8 mM EDTA). First, 30,000 cpm
of labeled DNA was incubated with increasing amounts of either YeMSPA or YeM-
YoeBSpn in 1× binding buffer, and the mixture was incubated at room temperature
for another 2 min. The reaction was then stopped by adding 14.7 µl of stop solution
(0.041 M thiourea, 1.5 M NaAc, pH 6, and 0.68 mg/ml tRNA) and 187 µl of absolute ethanol, followed by freezing at −70°C for 30 min. The mixture was then
centrifuged, and the pellet was washed with chilled 70% ethanol. The samples
were then treated as described in the DNase I footprinting assays. The
sequencing ladder prepared using the Maxam and Gilbert reaction for G + A (29)
was run along with the reaction samples.

RESULTS

Bioinformatics analysis of the S. pneumoniae-encoded YeM-
YoeBSpn locus. To determine whether the presence of the yeM-
YoeBSpn locus was universal in the genomes of pneumococcal
strains, BLASTP analyses were performed using the 22 anno-
tated pneumococcal genomes in the NCBI database. The re-

| Recombinant construct | E. coli host strain | Primer pair(s) | Description |
|-----------------------|---------------------|----------------|-------------|
| pQF_M1yMyB           | TOP10               | yeM7-F and yeM7-R | YeM_yoeBSpn along with 237 bp of its upstream sequence cloned into Ncol-BamHI site of pQF52, preceding the promoterless lacZ gene; using site-directed mutagenesis, two base pairs were substituted, which led to the 7th codon for YeMSPA, “S,” being replaced by an amber stop codon. |
| pQF_M2yMyB           | TOP10               | yeM273233-F and yeM273233-R | YeM_yoeBSpn along with 237 bp of its upstream sequence cloned into Ncol-BamHI site of pQF52, preceding the promoterless lacZ gene; using site-directed mutagenesis, four base pairs were substituted, which led to the 27th, 32nd, and 33rd codons for YeMSPA, “L,” “K,” and “N,” being replaced by amber, ochre, and amber stop codons, respectively. |
| pQF_M2yMyB           | TOP10               | yeM273233-F and yeM273233-R | YeMSPA-yoeBSpn along with 237 bp of its upstream sequence cloned into Ncol-BamHI site of pQF52, preceding the promoterless lacZ gene; using site-directed mutagenesis, two base pairs were substituted, which led to the 27th, 32nd and 33rd codons for YeMSPA, “L,” “K,” and “N,” being replaced by amber, ochre, and amber stop codons, respectively. |
| pLN_yM               | TOP10               | pLNyteM-F and pLNyteM-R | YeMSPA cloned into Ndel-I HindIII site of pLBAD, downstream of the PBAD promoter. |
| pLN_yMyB             | TOP10               | pLNyteM-F and pLNyteM-R | YeMSPA cloned into Ndel-I HindIII site of pLBAD, downstream of the PBAD promoter. |
| pEMH110              | BL21-CodonPlus(DE3)-RIL | pETyeM-F and pETyeM-B-R | YeMSPA-yoeBSpn start from GTG as annotated in S. pneumoniae R6, which is 36 bp upstream of ATG start site annotated in S. pneumoniae TIGR4, cloned into Ndel-BamHI site of pET28a, in frame with (His)6 at region of yeMSPA, corresponding to N-terminal region. |
Results showed that 7 strains that harbor solo YefM Spn antitoxin without its YoeB Spn toxin counterpart were all mutated or truncated at the 3′/H11032 terminus, whereas no solo YoeB Spn homologues were identified. Only 15 strains with both YefM Spn and YoeBSpn intact TA pairs were further analyzed (see Table S1 in the supplemental material). In all of them, the yefMSpn stop codon was separated by 3 nt from the start codon of yoeBSpn (see below). We were also able to resolve a discordant result, since YefMSpn was annotated in the NCBI database to initiate from different start codons among the S. pneumoniae strains.

We purified the YefMSpn protein and determined the sequence of the first 10 residues to be MEAVLYSTFR, demonstrating that the protein initiates unambiguously from the ATG codon as annotated in the TIGR4 strain (see Fig. S1). Disregarding the discrepancy in the annotation, the translated amino acid sequences of YefMSpn among the strains were completely identical except for the 77th residue of YefM Spn, which was either a Thr or an Ile (see Fig. S1). All strains shared 100% identity in the yoeBSpn translated amino acid sequences except strain JJA, which showed 96% identity (see Fig. S2). The start codon of yoeBSpn among all the strains was also identical, and the first 10 residues (MLLKFTEDAW) were confirmed by N-terminal sequencing (see Fig. S2). A sequence analysis of the DNA region upstream of the coding sequence of yefMSpn predicted the presence of two putative σ70-dependent promoters 30 nt apart, designated P$_{yefM1}$ and P$_{yefM2}$, which were experimentally shown to exist (see below). In addition, a putative ribosome-binding site (rbs) (5′-AcGAGG-3′; lowercase letters indicate deviations from the consensus sequence), located 7 nt upstream of the ATG start codon, was observed (Fig. 1A). The sequence of promoter P$_{yefM2}$ matched with those of other pneumococcal promoters (44) and showed an almost perfect consensus of the 35 (5′-cTGACA-3′) and 10 (5′-TATAA-3′) regions separated by a 17-nt spacer (Fig. 1A). In contrast, the 35 (5′-cTGACA-3′) and 10 (5′-TATAAa-3′) regions of the P$_{yefM1}$ promoter were more dissimilar to the consensus, and these promoter elements presented a spacer of 16 nt instead of 17 nt, indicating a suboptimal promoter sequence (Fig. 1A).

While examining the genetic structure around the yefM-yoeBSpn locus, we detected the presence of a BOX element (Fig. 1A) (coordinates 1567052 to 1567160; accession no AE007317) which was placed in the intergenic region between yefMSpn and the upstream gene in the pneumococcal R6 strain. The translated product of this upstream gene (coordinates

---

**FIG. 1.** The yefM-yoeBSpn locus harbors two promoters. (A) Nucleotide sequence of the pneumococcal yefM upstream region along with the first 71 nt of the yefM coding sequence. The P$_{yefM1}$ and P$_{yefM2}$ promoters with their respective −10 and −35 sequences are indicated within boxes, as well as the putative rbs. The incomplete palindrome sequence, termed “PS,” is also depicted, with the center of the palindrome indicated by the apex of a filled triangle. The boxA subelement is indicated in red, whereas the boxC subelement is indicated in blue. The start codon of yefMSpn is underlined. Asterisks denote the two transcriptional start sites of yefM-yoeBSpn as determined by primer extension. (B) Transcriptional start sites of the yefM-yoeBSpn operon as determined by primer extension. A yefM-near (N) or yefM-far (F) specific primer was used to anneal to the RNA samples prepared from S. pneumoniae R6. The primer was extended, yielding transcripts of 92 nt at 50°C and 115 nt at 55°C. The C, G, T, and A DNA sequencing ladders were obtained using the dideoxy-mediated chain termination sequencing method by annealing the yefM-near or yefM-far primer to pEMBH13 (34).
BOX elements are abundant repeated sequences which are placed apparently at random within intergenic regions of the pneumococcal genome. They are present only in *S. pneumoniae* and closely related species, such as *Streptococcus mitis*, and they have been proposed to affect the expression of their neighboring genes (27). Generally, a BOX element is formed by the juxtaposition of three subelements, designated boxA (59 bp), boxB (45 bp), and boxC (50 bp), with boxB placed between boxA and boxC (6, 27). In the BOX element that was found upstream of *yefM*-*yoeB*, boxB was absent, and boxA was located immediately upstream of boxC; this BOX element was thus designated boxAC. A closer inspection of these sequences revealed that the sequence 1567053 to 1567161 of the pneumococcal genome. They are present only in *S. pneumoniae* and closely related species, such as *Streptococcus mitis*, and they have been proposed to affect the expression of their neighboring genes (27). Generally, a BOX element is formed by the juxtaposition of three subelements, designated boxA (59 bp), boxB (45 bp), and boxC (50 bp), with boxB placed between boxA and boxC (6, 27). In the BOX element that was found upstream of *yefM*-*yoeB*, boxB was absent, and boxA was located immediately upstream of boxC; this BOX element was thus designated boxAC. A closer inspection of these sequences revealed that the sequence 1567053 to 1567161 of the genome of *S. pneumoniae* was obtained by RT-PCR with total RNA isolated from *S. pneumoniae* R6 as a template and oligonucleotide GyoeBC as the primer. PCR was then carried out using the yefMNHis and yefMC primer pairs in which the RT reaction was omitted. Lane C, DNA molecular weight standard (Smartladder; Euregentec) with sizes as indicated.

![Diagram](http://jb.asm.org/Downloadedfrom)

**FIG. 2.** The *yefM*-*yoeB* locus is organized as an operon. (A) Genetic structure of the region. There are two *yefM* TGA stop codons (the first one highlighted in green) between the putative *yoeB* ribosome-binding site (GAGGAG; designated rbs YoeBSpn) and the ATG initiation codon of *yoeB*. The primers used in the RT-PCR or PCR are indicated (lines with their arrowheads pointing in the 5′ → 3′ direction) and designated yefMNHis and yefMC for *yefM* or GyoeBN and GyoeBC for *yoeB*. The PCR fragments and their expected sizes (336 bp for *yefM*, 300 bp for *yoeB*, and 563 bp for *yefM*-*yoeB*) are shown below. (B) RT-PCR results. A cDNA fragment spanning the *yefM*-*yoeB* genes was synthesized using total RNA prepared from *S. pneumoniae* R6 as a template and oligonucleotide GyoeBC as the primer. PCR was then carried out using the yefMNHis and yefMC primer pair on the cDNA synthesized (lane 1) and the chromosomal DNA of *S. pneumoniae* R6 (lane 2), the primer pair GyoeBN and GyoeBC on cDNA (lane 3) and chromosomal DNA (lane 4), and the primer pair yefMNHis and GyoeBC on cDNA (lane 5) and chromosomal DNA (lane 7). Lane 6, product of the PCR using the yefMNHis and GyoeBC primer pairs in which the RT reaction was omitted. Lane C, DNA molecular weight standard (Smartladder; Euregentec) with sizes as indicated.
out possible DNA contamination (Fig. 2B, lane 6). The resulting DNA products were of the expected sizes for yefM<sub>Spn</sub>, (336 bp; lanes 1 and 2), yoeB<sub>Spn</sub> (300 bp; lanes 3 and 4), and yefM-yoeB<sub>Spn</sub> (563 bp; lanes 5 and 7) (Fig. 2B). The latter cDNA corresponded to a single mRNA encompassing both genes, thus demonstrating that in <i>S. pneumoniae</i>, yefM<sub>Spn</sub> and yoeB<sub>Spn</sub> are cotranscribed. Similar results were obtained when RT-PCR was conducted on total RNA isolated from <i>E. coli</i> harboring a pGEM-T Easy recombinant plasmid which contained the entire yefM-yoeB<sub>Spn</sub> reading frames along with 237 bp of sequences upstream of yefM<sub>Spn</sub> encompassing the putative promoter sequences (not shown). These results are in agreement with those reported for other TA loci, such as relBE2 in <i>S. pneumoniae</i> (35) and the kis-kid genes of the <i>pamD</i> maintenance system of plasmid R1 (43), where the antitoxin and toxin genes are cotranscribed, which is also a general characteristic of TAS.

The yefM-yoeB<sub>Spn</sub> operon is transcribed from two promoters, P<sub>yefM1</sub> and P<sub>yefM2</sub>. To determine the transcription initiation start site(s) of the yefM-yoeB<sub>Spn</sub> operon, primer extension analyses were performed. Total RNA was prepared from exponentially growing <i>S. pneumoniae</i> R6 cells, annealed with two labeled primers, termed yefM-near and yefM-far, and extended with reverse transcriptase at either 50°C or at 55°C. Two different extension products, of 92 nt or 115 nt, were observed (Fig. 1B), which indicated that the yefM-yoeB<sub>Spn</sub> operon was transcribed from two transcriptional start sites - the “A” purine residues located 84 nt and 25 nt upstream of the yefM<sub>Spn</sub> ATG start codon (see Fig. 1A). An inspection of the DNA sequence around the transcriptional start sites revealed that the primer extension products would be complementary to mRNAs transcribed from the P<sub>yefM1</sub> and P<sub>yefM2</sub> promoters, respectively. Our primer extension results showed that the transcriptional start sites for both P<sub>yefM1</sub> and P<sub>yefM2</sub> are located 5 nt downstream relative to their respective -10 elements (Fig. 1A). We can conclude that insertion of the boxAC element had indeed led to the creation of a functional P<sub>yefM1</sub> promoter.

**Determination of promoter activities in vivo.** The in vivo functionality and strength of the two promoters were determined by employment of transcriptional fusions with a promoterless lacZ gene as a reporter in plasmid pQF52 in <i>E. coli</i>. The more important results obtained are depicted in Fig. 3. The functionality of the promoters was validated when <i>E. coli</i> cells harboring P<sub>yefM1</sub> (pQF-P2) showed a mean ß-galactosidase activity of 122 ± 36 Miller units (MU) (Fig. 3). The ß-galactosidase activities of the other constructs were stated as ratios by normalizing with those reported for other TA loci, such as the parD activity in <i>E. coli</i> (47), where the antitoxin and toxin genes are cotranscribed, which is also a general characteristic of TAS. The yefM<sub>Spn</sub> antitoxin was expressed in <i>E. coli</i> (Fig. 1B), which indicated that the yefM<sub>Spn</sub>-galactosidase activity was decreased by 7% compared to that of the P<sub>yefM1</sub>-galactosidase activity (i.e., a ratio of 1.89) was detected in the mutated clone, indicating that the increased activity was still present (Fig. 3). Even when yoeB<sub>Spn</sub> was included along with this mutated construct, the activation was still evident, with an observed ratio of 2.03 (Fig. 3). No detectable ß-galactosidase activity was observed in cells that harbored just the yefM<sub>Spn</sub>-reading frame (Fig. 3). Taken together (including results obtained from the other recombinants in Table 1), these results suggested that the increased ß-galactosidase activity was observed when the yefM<sub>Spn</sub> coding sequence was present along with the entire upstream region spanning boxAC, P<sub>yefM1</sub>, and P<sub>yefM2</sub> in cis. Furthermore, activation was not due to YefM<sub>Spn</sub> acting as an activator or to the presence of an internal promoter within the yefM<sub>Spn</sub> reading frame.

Unexpectedly, a functional promoter was detected within the yoeB<sub>Spn</sub>-coding sequence, and this promoter was likely present at the C terminus (the last 159 nt of yoeB<sub>Spn</sub> in pQF-CyB gave an activity ratio of 1.28) (Table 1). To verify whether this promoter was indeed functional in its native host, a 159-nt fragment encoding the C terminus of YoeB<sub>Spn</sub> was cloned into the promoter-probe PAST plasmid, which carries the green fluorescent protein-encoding gene as a reporter (42); transformants were rescued in <i>S. pneumoniae</i> R6. Two clones having the insert in opposite directions were tested for fluorescence, and no promoter activity was evident (not shown). The genome of <i>S. pneumoniae</i> has a high A+T content (61%) (5); thus, sequences resembling a –10 region are likely to exist by chance, and they could be recognized by the <i>E. coli</i> α<sup>70</sup> RNA polymerase (RNAP) (47).

**Determination of DNA binding sites of the YefM<sub>Spn</sub> protein and the YefM-YoeB<sub>Spn</sub> protein complex.** EMSA and footprint-
ing experiments were carried out to determine the DNA binding sites for the YefM<sub>Spn</sub> antitoxin and the YefM-YoeB<sub>Spn</sub> TA complex. For EMSA, a 322-bp [γ-<sup>32</sup>P]ATP-labeled PCR-amplified DNA fragment encompassing the entire upstream region of yefM<sub>Spn</sub>, along with the boxAC element and the PS, was incubated with increasing amounts of purified YefM<sub>Spn</sub> or purified YefM<sub>Spn</sub>-YoeB<sub>Spn</sub> proteins. The unbound DNA fragments and the nucleoprotein complexes were observed as well-separated bands on native 5% polyacrylamide gels, indicating that the antitoxin alone or the protein complex was bound to this DNA fragment. Furthermore, preliminary results indicated that both proteins were bound to a DNA fragment containing only PS, and no other binding site was observed (not shown). Consequently, a 284-bp DNA fragment encompassing PS (but not the boxAC element) was used for further assays. This DNA fragment was end labeled with [γ-<sup>32</sup>P]ATP on the coding strand. Aliquots (3,000 cpm) were incubated with increasing amounts of either purified YefM<sub>Spn</sub> or YefM<sub>Spn</sub>-YoeB<sub>Spn</sub> from the P<sub>BAD</sub> promoter. β-Galactosidase activity levels are the average results of eight independent experiments and are presented as relative levels (ratios) by normalization of the β-galactosidase activity levels with that of pQF_P2, which is 122 ± 36 MU (a detailed presentation of all the constructs and their resulting β-galactosidase activity levels is depicted in Table 1).
PS palindrome sequence upstream of yefM-yoeBSpn labeled 284-bp DNA fragment (3,000 cpm per lane) encompassing the gels. The unbound DNA fragments and the nucleoprotein complexes were separated by electrophoresis in 5% native polyacrylamide gel. The unbound DNA fragments and the nucleoprotein complexes were separated by electrophoresis in 5% native polyacrylamide gel.

![Image](http://jb.asm.org/)

**FIG. 4.** EMSA showing in vitro binding of the purified YefMSpn protein and the YefM-YoeBSpn protein complex with a [γ-32P]ATP-labeled 284-bp DNA fragment (3,000 cpm per lane) encompassing the PS palindrome sequence upstream of yefM-yoeBSpn with 10 ng/µl heparin as a competitor. Increasing amounts of YefMSpn (from left to right, 0.00, 1.00, 1.50, 2.00, 2.50, 3.00, 3.50, 4.00, and 5.00 µg) (A) or YefM-YoeBSpn (from left to right, 0.00, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.40, and 0.50 µg) (B) were incubated with the labeled DNA fragment. The unbound DNA fragments and the nucleoprotein complexes were separated by electrophoresis in 5% native polyacrylamide gels.

Refinement of the contacts between the YefMSpn and YefM-YoeBSpn proteins with their DNA target was achieved by hydroxyl radical footprinting because of the high resolution potential of this technique (52). All of the DNase I footprints were also detectable by hydroxyl radical footprinting, although the former were divided in two, thus defining a region of palindromic symmetry in which two and three protected bases on the coding strand were mirrored by three and two protected bases on the noncoding strand (Fig. 5 and 6). A very weak footprint was observed when YefMSpn alone was used (not shown), corroborating the low affinity of the antitoxin for its target. Again, only the results obtained with the TA complex are presented, and a summary of the footprints obtained with the YefMSpn antitoxin alone is depicted in Fig. S4 in the supplemental material. The protections by YefM-YoeBSpn in the hydroxyl radical assays (Fig. 5C) were included within those observed for the DNase I footprinting. On the coding strand (Fig. 5C), the two protected regions were separated by 9 nt; these were located at nucleotides −11 and −10 of the left arm, as well as from nucleotide +1 to +3 of the right arm of PS, which overlapped with the −35 element of P_{yefM2} (Fig. 5C). For the noncoding strand, the protected sites spanned from nucleotides −3 to −1 of the left arm and at nucleotides +9′ and +10′ of the right arm of PS (Fig. 5C). An overall picture of the footprinting analyses on this DNA region (Fig. 6A) could be better visualized when the protected regions were depicted on a schematic drawing of a 10.5-bp DNA helix (Fig. 6B). No indication of protein-induced DNA bends was observed; however, three hyperepoxified bases were detected on the coding strand at the border of the footprint located 3′ to the −35 promoter region (Fig. 6A), indicative of a local helix deformation at these positions.

**DISCUSSION**

We show here that the pneumococcal yefM-yoeBSpn TAS is constituted as an operon in which two promoters direct the synthesis of the mRNA transcript. The genetic regulation of the pneumococcal yefM-yoeBSpn locus differed from most conventional TA loci due to the presence of the two promoters, P_{yefM1} and P_{yefM2}, and the boxAC element. Promoter P_{yefM2}, which is closer to the yefM-yoeBSpn locus, is likely the "original" or native promoter of the TA locus, whereas P_{yefM1}, which is the weaker of the two promoters, came about from the insertion of boxAC upstream of P_{yefM2}. Thus, the presence of this BOX element provides a unique scenario for regulation of the expression of this particular pneumococcal TAS. If the regulation of the yefM-yoeBSpn locus is considered in terms of the P_{yefM2} Promoter alone, then the locus is negatively auto-regulated like other typical TAS. The YefMSpn antitoxin would

![Image](http://jb.asm.org/)
repress, albeit weakly, transcription from P<sub>yefM2</sub>, whereas the YoeB<sub>Spn</sub> toxin would act as a corepressor to fully regulate transcription from P<sub>yefM2</sub>. Both DNase I and hydroxyl radical footprinting assays showed that YoeB<sub>Spn</sub> increased the affinity of YefM<sub>Spn</sub> for the operator site. The 35 region of P<sub>yefM2</sub> was not directly protected by the YefM<sub>Spn</sub> protein (see Fig. S4 in the supplemental material), but protection of this region was very clear in the presence of the YefM-YoeB<sub>Spn</sub> complex (Fig. 6). Representation of the distribution of bases which were protected by bound YefM-YoeB<sub>Spn</sub> proteins (Fig. 6) on a DNA molecule with a helical periodicity of 10.5 bp per helix turn (41) clearly showed that the proteins contact the DNA backbone on one face of the helix and that the protections did not extend further than the central region of the PS. The DNase I footprints delineated two central regions of 10 and 19 protected bases (on the coding strand) separated by two unprotected ones; on the noncoding strand, a single footprint could be observed. However, the high-resolution analyses ob-

FIG. 5. DNase I and hydroxyl radical footprinting assays for the purified YefM-YoeB<sub>Spn</sub> proteins on the DNA fragment containing the PS palindrome sequence. (A and B) DNase I footprinting assays. Both coding strands (A) and noncoding strands (B) of [γ<sup>32</sup>P]ATP-labeled DNA fragments containing the PS (30,000 cpm) were incubated with increasing amounts of YefM-YoeB<sub>Spn</sub> protein complex (0, 1, 2, and 5 μg) prior to DNase I digestion. The reaction mixture was then separated on a 8% polyacrylamide gel containing 7 M urea along with the DNA sequencing ladder prepared using the dideoxy sequencing reaction. Nucleotide sequences protected by the YefM-YoeB<sub>Spn</sub> complex from DNase I digestion (green) and hypersensitivity observed at these nucleotides (brown) are indicated. (C) Hydroxyl radical footprinting assays. Aliquots (30,000 cpm) of both coding (lanes 1 and 2) and noncoding (lanes 3 and 4) strands of a [γ<sup>32</sup>P]ATP-labeled DNA fragment containing the PS were incubated with (lanes 2 and 3) or without (lanes 1 and 4) 5 μg of purified YefM-YoeB<sub>Spn</sub> protein complex prior to hydroxyl radical treatment. The reaction mixture was then separated on a 8% polyacrylamide gel containing 7 M urea along with a DNA sequencing ladder (A/G) prepared by the Maxam-Gilbert sequencing method. Nucleotide sequences of the coding and noncoding DNA strands of the yefM<sub>Spn</sub> upstream region containing the PS that were protected from hydroxyl radical attack by the YefM-YoeB<sub>Spn</sub> proteins are shown in red.
tained with the hydroxyl radical footprinting showed that the DNase I footprints could be further separated into two protected regions on both coding and noncoding strands (Fig. 6). On the basis of these results, it would appear that the mechanism of transcriptional repression from \( P_{\text{yefM}1} \) is likely due to the binding of the antitoxin or the TA complex to the PS site, which overlaps the \(-35\) region. Occupancy of the TA complex on the PS would lead to interference of the activity of RNAP by a blue bar.

FIG. 6. Nucleotide sequence of the coding and noncoding DNA strands of the \( yefM_{\text{Spn}} \) upstream region containing the PS palindrome sequence. (A) The \(-10\) and \(-35\) regions of the \( P_{\text{yefM2}} \) promoter are indicated. The PS is indicated by arrows, and the center of the PS is depicted with the apex of a filled triangle. The nucleotides protected from DNase I digestion by the YefM-YoeB\( \text{Spn} \) proteins are highlighted in green. The nucleotides protected from hydroxyl radical attack by the YefM-YoeB\( \text{Spn} \) proteins are indicated by red arrows. The nucleotides with hypersensitivity are indicated in brown. (B) Representation of the double helix of a B-DNA fragment encompassing the \( yefM_{\text{Spn}} \) upstream region containing the PS. Nucleotides protected by the YefM-YoeB\( \text{Spn} \) protein complex from DNase I digestion (green circles) and hydroxyl radical attack (highlighted in red) are indicated. The nucleotides with hypersensitivity are also indicated (brown circles). The \(-35\) region of the \( P_{\text{yefM2}} \) promoter is represented by a blue bar.

Intriguingly, even though the overall promoter activity of both \( P_{\text{yefM1}} \) and \( P_{\text{yefM2}} \) was lower than that of the \( P_{\text{yefM2}} \) promoter alone, the overall promoter activity increased dramatically when the \( yefM_{\text{Spn}} \) reading frame was provided in cis. However, transcriptional activation by the YefM\( \text{Spn} \) protein was ruled out, since this activation was not observed when \( yefM_{\text{Spn}} \) was provided in trans and, more tellingly, stop codons introduced into the \( yefM_{\text{Spn}} \) reading frame did not abrogate transcriptional activation. No internal promoters were detected within the \( yefM_{\text{Spn}} \) reading frame. This activation was likely due to cis-acting elements, and perhaps the boxAC element and/or its downstream sequences contain activator binding sites for a host factor(s) that has yet to be elucidated.

BOX elements have been associated with the following: (i) genes involved in the induction of pneumococcal genetic competence (17, 27), (ii) virulence-related genes (27), and (iii) expression of phase variation (45). Another scenario was found for the \( S.\ pneumoniae \) maltose regulon, where a full BOX element was detected upstream of the regulatory \( malAR \) genes, indicating that the role of these repeated elements may be more complicated than envisaged (40). However, since a large part of the pneumococcal genome is devoted to sugar uptake and metabolism, it seems possible that regulation of these processes (metabolism and response to stress) is important for the bacterial fitness in its natural niche (the human nasopharynx) and may play a role during invasion of other tissues, such as the lungs (14).

In the case of the pneumococcal \( yefM-yoeB_{\text{Spn}} \) TAS, the presence of the boxAC element also added another level of complexity to the regulation of this operon. The two promoters driving the \( yefM-yoeB_{\text{Spn}} \) locus may thus enable a more dynamic regulation of the TA operon. Since \( P_{\text{yefM1}} \) does not appear to be regulated by either YefM\( \text{Spn} \) or the YefM-
YoeBSpn complex, P_{yefM} would provide a higher basal level of transcription for the yefM-yoeBSpn operon, thus enabling cells to have a faster response to any sudden changes in their environment.

BOX elements may not serve a specialized role in the host; however, their presence is more likely to affect the regulation or expression of the genes by providing a versatile response to environmental changes. Be that as it may, from an evolutionary point of view (2), the insertion of a translocase piece (the boxAC element) into an operative piece (the environmental changes. Box elements modulate gene expression in Streptococcus pneumoniae; impact on the fine-tuning of competence development. J. Bacteriol. 188:879–892.

30. Makarova, K., Y. I. Wolf, and E. V. Koonin. 2003. Comprehensive comparison-genome analysis of type II toxin-antitoxin systems and related mobile stress response systems in prokaryotes. Biol. Direct 4:19.

31. Mandell, L. A., et al. 2007. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. Clin. Infect. Dis. Suppl. 2:257–S72.

32. Marianoisky, L., E. Aizenman, H. Engelberg-Kulka, and G. Glaser. 2001. The regulation of the Escherichia coli mazEF promoter involves an unusual palindromic motif. J. Bacteriol. 183:5975–5984.

33. Martin, B., et al. 1992. A highly conserved repeated DNA element located in the chromosome of Streptococcus pneumoniae. Nucleic Acids Res. 20:3479–3483.

34. Masuda, Y., K. Miyakawa, Y. Nishimura, and E. Ohtsubo. 1993. chpA and chpB, Escherichia coli chromosomal homologs of the pmrA locus responsible for stable maintenance of plasmid R100. J. Bacteriol. 175:6850–6856.

35. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499–560.

36. McLean, B. W., S. L. Wiseman, and A. M. Kropinski. 1997. Functional analysis of sigma-70 consensus promoters in Pseudomonas aeruginosa and Escherichia coli. Can. J. Microbiol. 43:981–985.

37. Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

38. Nagawa, F., et al. 1998. Functional analysis of the RAG protein recombination signal sequence complex for V(D)J type recombination. Mol. Cell. Biol. 18:655–663.

39. Nesin, M., J. R. Lupski, and G. N. Godson. 1988. Role of the 5’ upstream sequence and tandem of conserved 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. Biol. Direct 4:19.

40. Nieto, C., et al. 2007. The yefM-yoeB toxin-antitoxin systems of Escherichia coli and Streptococcus pneumoniae: functional and structural correlation. J. Bacteriol. 199:1266–1278.

41. Nieto, C., et al. 2006. The chromosomal relBE toxin-antitoxin locus of Streptococcus pneumoniae: characterization and use of a bioluminescence resonance energy transfer assay to detect toxin-antitoxin interaction. Mol. Microbiol. 59:1260–1269.

42. Nieto, C., E. Sadowsky, A. G. de la Campa, W. Hryniewicz, and M. Espinosa. 2010. The relBE2spn toxin-antitoxin system of Streptococcus pneumoniae: role in antibiotic tolerance and functional conservation in clinical isolates. PLoS ONE 5:e11289.

43. Pandey, D. P., and K. Gerdes. 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucleic Acids Res. 33:966–976.

44. Pedersen, K., K. S. Christensen, and K. Gerdes. 2002. Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. Mol. Microbiol. 45:501–510.

45. Puyet, A., and M. Espinosa. 1993. Structure of the multilactein-uptake locus of Streptococcus pneumoniae: correlation to the Escherichia coli maltose regulon. J. Bacteriol. 175:2804–2811.

46. Rhodes, D., and A. Klug. 1980. Helical periodicity of DNA determined by enzyme digestion. Nature 286:573–578.
42. Ruiz-Cruz, S., V. Solano-Collado, M. Espinosa, and A. Bravo. 2010. Novel plasmid-based genetic tools for the study of promoters and terminators in Streptococcus pneumoniae and Enterococcus faecalis. J. Microbiol. Methods 83:156–163.
43. Ruiz-Echevarría, M. J., A. Berzal-Herranz, K. Gerdes, and R. Díaz-Orejas. 1991. The kis and kid genes of the parD maintenance system of plasmid R1 form an operon that is autoregulated at the level of transcription by the co-ordinated action of the Kis and Kid proteins. Mol. Microbiol. 5:2685–2693.
44. Sabelnikov, A. G., B. Greenberg, and S. A. Lacks. 1995. An extended /H11002 10 promoter alone directs transcription of the DpnII operon of Streptococcus pneumoniae. J. Mol. Biol. 250:144–155.
45. Saluja, S. K., and J. N. Weiser. 1995. The genetic basis of colony opacity in Streptococcus pneumoniae: evidence for the effect of box elements on the frequency of phenotypic variation. Mol. Microbiol. 16:215–227.
46. Sambrook, J., and D. W. Russell. 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
47. Samul, R., and F. Leng. 2007. Transcription-coupled hypernegative supercoiling of plasmid DNA by T7 RNA polymerase in Escherichia coli topoisomerase I-deficient strains. J. Mol. Biol. 374:925–935.
48. Shao, Y., et al. 2011. TADB: a web-based resource for type 2 toxin-antitoxin loci in bacteria and archaea. Nucleic Acids Res. 39:D606–D611.
49. Tettelin, H., et al. 2001. Complete genome sequence of a virulent isolate of Streptococcus pneumoniae. Science 293:498–506.
50. Tian, Q.-B., et al. 2001. Specific protein-DNA and protein-protein interaction in the hig gene system, a plasmid-borne proteic killer gene system of plasmid Rts1. Plasmid 45:63–74.
51. Tsuchimoto, S., and E. Ohtsubo. 1993. Autoregulation by cooperative binding of the PemN and PemK proteins to the promoter region of the pem operon. Mol. Gen. Genet. 237:81–88.
52. Tullius, T. D., and B. A. Dombroski. 1986. Hydroxyl radical “footprinting”: high resolution information about DNA-protein contacts and application to lambda repressor and Cro proteins. Proc. Natl. Acad. Sci. U. S. A. 83:5469–5473.
53. Van Melderen, L. 2010. Toxin-antitoxin systems: why so many, what for? Curr. Opin. Microbiol. 13:781–785.
54. Van Melderen, L., and M. Saavedra De Rast. 2009. Bacterial toxin-antitoxin systems: more than selfish entities? PLoS Genet. 5:e1000437.
55. Weihofen, W. A., A. Cicek, F. Pratto, J. C. Alonso, and W. Saenger. 2006. Structures of Omega repressors bound to direct and inverted DNA repeats explain modulation of transcription. Nucleic Acids Res. 34:1450–1458.
56. Zhu, L., et al. 2006. Characterization of mRNA interferases from Mycobacterium tuberculosis. J. Biol. Chem. 281:18638–18643.