JOURNAL OF BACTERIOLOGY, Sept. 1994, p. 5820–5830
0021-9193/94/$04.00+0

Vol. 176, No. 18

Genetic Control of Bacterial Suicide: Regulation of the Induction of PBSX in *Bacillus subtilis*

GERALD E. McDONNELL,† HEATHER WOOD,‡ KEVIN M. DEVINE, AND DAVID J. McCONNELL*

Department of Genetics, Trinity College, Dublin, Ireland

Received 28 September 1993/Accepted 30 June 1994

PBSX is a phage-like bacteriocin (phibacin) of *Bacillus subtilis* 168. Bacteria carrying the PBSX genome are induced by DNA-damaging agents to lyse and produce PBSX particles. The particles cannot propagate the PBSX genome. The particles produced by this suicidal response kill strains nonlysogenic for PBSX. A 5.2-kb region which controls the induction of PBSX has been sequenced. The genes identified include the previously identified repressor gene *xre* and a positive control factor gene, *pcf*. Pcf is similar to known sigma factors and acts at the late promoter P_L, which has been located distal to *pcf*. The first two genes expressed from the late promoter show homology to genes encoding the subunits of phage terminases.

The defective bacteriophage PBSX of *Bacillus subtilis* 168 is a phage-like bacteriocin, or phibacin, with curious biological properties. PBSX lysogens exposed to DNA-damaging agents produce PBSX phage-like particles which kill other nonlysogenic *Bacillus* strains (25) by binding to and disrupting the cell wall. Each PBSX particle has a small head and a relatively long tail. Phage heads appear to contain randomly selected 13-kb segments of host DNA, which are not injected into susceptible cells. PBSX is therefore not propagated by these particles (25, 26). PBSX appears to be a particulate bacteriocin which has evolved from a bacteriophage.

Induction of PBSX is controlled by the repressor gene *xre* (6, 27, 40, 41). The amino acid sequence of Xre predicts that it is a helix-turn-helix (HLH) protein resembling other DNA-binding proteins such as lambda cI and Cro (14, 41). However, little is known about the mechanism of action of Xre, and apart from the suggestion that *xre* and a neighboring gene, now called open reading frame 10 (ORF10), are regulated by Xre itself, nothing is known about how the late genes are regulated (41).

All known structural and lytic protein genes have been shown to be clustered within a large operon of at least 19 kb in length, which appears to be expressed from a single late promoter region called P_L (Fig. 1a) (40). A number of mutations in regulatory genes, including *xin* (noninducible for PBSX) and *xhi* (heat inducible for PBSX), are located proximal to mutations affecting phage head and tail proteins such as *xhd*, *xtl*, and *xki* (6, 38). A regulatory mutation, *xhi-1479*, renders PBSX thermoinducible (6). A 1.2-kb *Eco*RI fragment from a PBSX wild-type strain was found to complement the *xhi-1479* mutation, and sequence analysis identified the *xre* gene as the site of the mutation (41). *xhi-1479* is thus analogous to the *cIts857* mutation of bacteriophage lambda (36). An insertion mutation called 316, which lies upstream of the major late operon, abolishes induction of the prophage (40). This finding suggests that there are other genes, possibly between *xre* and the late operon, which are required for late gene expression.

In this report, we present the sequence and analysis of a 5.2-kb segment of DNA which spans the repressor gene *xre*, the site of insertion directed by fragment 316, and the late promoter P_L. A model which describes the roles of these genes and sites in the regulation of PBSX late gene expression is presented.

MATERIALS AND METHODS

Bacterial strains and plasmids. Bacterial strains and plasmids are listed in Table 1.

Media. *B. subtilis* and *Escherichia coli* were grown in Luria-Bertani (LB) broth (1% tryptone, 0.5% yeast extract, 0.5% NaCl) or LB agar. For selection, media contained chloramphenicol (at 3 μg/ml for low-copy-number vectors and 5 μg/ml for higher-copy-number vectors) and kanamycin (5 μg/ml) for *B. subtilis*, or chloramphenicol (15 μg/ml) and ampicillin (100 μg/ml) for *E. coli* as appropriate. α-Amylase activity was detected by adding starch (0.2%) to the media, and plates were stained over iodine crystals. For selection of auxotrophic markers in *B. subtilis*, cells were grown in 1× Spizizen salts [0.6% KH_2PO_4, 1.4% $KHPO_4$, 0.2% $(NH_4)_2SO_4$, 1.1% $Na_3C_6H_5O_7 \cdot 2H_2O$, 0.02% $MgSO_4 \cdot 7H_2O$] supplemented with 0.5% glucose, appropriate amino acids (0.005%), and nucleosides (0.01%). PBSX was induced by mitomycin at a final concentration of 0.1 μg/ml (27).

Transformation. *B. subtilis* competent cells were prepared and transformed as described by Contente and Dubnau (10). *E. coli* competent cells were prepared and transformed as described by Cohen et al. (9).

DNA preparations. All DNA manipulations for both *E. coli* and *B. subtilis* were performed as described by Sambrook et al. (30). *B. subtilis* chromosomal DNA was prepared by the method of Rodriguez and Tait (29).

RNA preparation. *B. subtilis* RNA was isolated as described by O'Reilly et al. (28). Cells were filtered through 115-ml Nalgene filters (0.45-mm pore size), washed with 1 ml of STET buffer (8% sucrose, 0.5% Triton X-100, 50 mM Tris-Cl [pH 8.0], 10 mM EDTA, 4 mg of lysozyme per ml), and incubated on ice for 15 min to allow for cell lysis. The lysate was extracted twice with 2 ml of phenol-chloroform (1:1) and once with chloroform. RNA was precipitated with 500 μl of 7.5 M ammonium acetate and 5 ml of ethanol, spun at 10,000 × *g* for 30 min (at 4°C), and washed with 70% ethanol. Preparations were resuspended in H_2O and adjusted to a concentration of 2

* Corresponding author. Mailing address: Department of Genetics, Lincoln Place Gate, Trinity College, Dublin 2, Ireland.

† Present address: Department of Microbiology, Colorado State University, Fort Collins, CO 80523.

‡ Present address: European Molecular Biology Laboratory, 6900 Heidelberg, Germany.

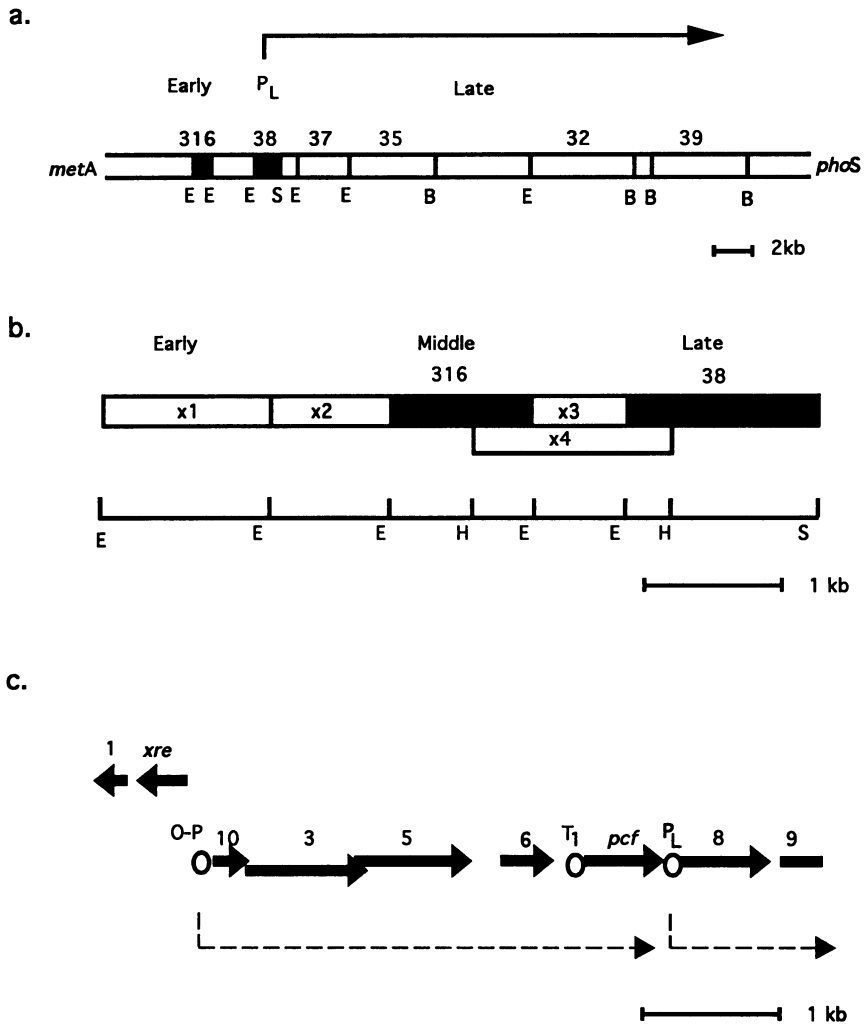


FIG. 1. (a) Structural map of the PBSX genome, as determined by Wood et al. (40). PBSX maps between *metA* and *pho*A on the genetic map of *B. subtilis* 168. Fragments previously used by Wood et al. (40) to identify PBSX operons are indicated at the top. The early region was defined by integration of fragment 316, which abolishes PBSX induction. The late operon encodes phage structural and lytic proteins, which are expressed from a single promoter (P_L) within fragment 38. Restriction sites: B, *Bam*HI; E, *Eco*RI; S, *Sac*I. (b) Structural map of the early and middle regions of PBSX. Fragments 316 and 38, as discussed above, are shown in boldface. Fragment x1 complements the *xhi-1479* mutation, which renders PBSX thermoinducible, and encodes a repressor protein (Xre) of PBSX. The orientation of fragments x1, x2, 316, x4, and 38 was determined by Southern analysis (results not shown). The early, middle, and late regions are defined in the text. Restriction sites: E, *Eco*RI; H, *Hin*dIII; S, *Sac*I. (c) ORFs identified in the 5.2-kb region of PBSX. Six complete ORFs are in the same orientation, while *xre* is in the opposite orientation. It is proposed that Xre is autoregulated by selective binding to four operator sites within the operator-promoter (O-P) region. We propose that upon induction of PBSX, ORF10, ORF3, ORF5, ORF6, and ORF7 are expressed as a single transcriptional unit, as indicated by dotted lines. T_1 is similar to known rho-independant transcriptional terminators, suggesting that transcription may terminate or pause at this site. Antitermination, possibly directed by ORF5, would allow expression of Pcf and therefore transcription from P_L into the late genes. The first two of these late genes, ORF8 and ORF9, appear to encode the small and large subunits of a terminase.

μg/μl. All solutions were treated with 0.1% diethylpyrocarbonate.

RNA samples were treated with DNase (Promega, Madison, Wis.) in 0.4 M Tris-Cl (pH 8.0)–60 mM $MgCl_2$ for 1 h at 37°C, extracted with phenol-chloroform, and ethanol precipitated.

Unidirectional deletions. Unidirectional deletions were prepared by using an Erase-a-Base kit (Promega). Plasmid DNA (5 to 10 μg) was digested with two restriction enzymes to give a protected end with a 3′ overhang and an end with a 5′ overhang from which exonuclease III can delete. Digestions were extracted with phenol-chloroform and ethanol precipitated. DNA was resuspended in 62 μl of 1× exonuclease III buffer (66 mM Tris-Cl [pH 8.0], 0.66 mM $MgCl_2$) and heated to 30°C; 300 U of exonuclease III was added, and 2.5-μl aliquots were removed at 30-s intervals into 7.5 μl of SI mix (172 μl of water, 27 μl of 7.4× SI buffer [0.3 M potassium acetate {pH 4.6}; 2.5 M NaCl, 10 mM $ZnSO_4$, 50% glycerol], 60 U of S1 nuclease). Reactions were left at 20°C for 30 min, stopped by addition of 1 μl of SI stop (0.3 M Tris, 100 mM EDTA), and ethanol precipitated. Deletions were treated with Klenow enzyme and religated as previously described.

Radioactive labelling of DNA. Oligonucleotides were end labelled by using T4 polynucleotide kinase enzyme and 10 pmol of oligonucleotide in 600 mM Tris-Cl (pH 9.5)–100 mM

TABLE 1. Bacterial strains and plasmids used

Plasmid or strain	Description	Reference or source
Plasmid		
pHV1435h	EMBL3 λ clone C, overlapping early region of PBSX, cloned into *SalI* of pHV1435	40
pWD38	Ap^r Cm^r fragment 38 fused to promoterless *amyL*	40
pEB113	Kn^r Ap^r	25
pDG268	Ap^r Cm^r, promoterless *lacZ*	3
E. coli TGI	K-12 Δ(*lac-pro*) *supE thi hsdR* F′ *traD36 proAB lacI lacZ* ΔM15	Amersham
B. subtilis		
IA420	*ilvA1 metB5 xhi-1479 purA16 xki-1479*	6
RB1081	*pro(AB) pyrX* ΔPBSX	7

$MgCl_2$–1 mM EDTA–10 mM spermidine–50 mM dithiothreitol with [γ-^{32}P]dATP (3,000 Ci mmol^{-1}). Reactions were carried out at 37°C for 60 min, and T4 polynucleotide kinase was denatured at 65°C for 10 min.

Primer extension. Primer extension analysis was done as described by O'Reilly et al. (28). Twenty picomoles of primer was labelled as described above with T4 polynucleotide kinase in 10 μl. Two microliters of labelled primer was annealed to 30 μg of total RNA in 1× hybridization buffer (0.1 M KCl, 0.05 M Tris-Cl [pH 8.3]) in a final volume of 10 μl. Annealing was carried out at 90°C for 1 min, at 60°C for 2 min, and on ice for 15 min. Seven microliters of the annealing mix was added to 2 μl of 5× reverse transcriptase buffer (250 mM Tris-Cl [pH 8.3], 200 mM KCl, 36 mM magnesium acetate, 1 mM each deoxynucleoside triphosphate [dNTP], 2 U of RNasin [an RNase inhibitor from Promega] per ml) and 1 μl (2 U) of avian myeloblastosis virus reverse transcriptase (Promega). Reaction mixtures were incubated at 50°C for 1 h. DNA was precipitated with ethanol, resuspended in sequencing loading dye (U.S. Biochemical [Cleveland, Ohio] sequencing kit), boiled for 1 min, and loaded onto a 6% acrylamide gel (running in 1× Tris borate buffer), in parallel with sequencing reactions.

DNA sequencing. The sequence of nucleotides was determined by the dideoxy-chain termination method essentially as described by Sanger et al. (31). Double-stranded DNA was isolated from 10 ml of boiling preparations and treated with 2 U of RNase at 50°C for 15 min. DNA was denatured in 0.2 M NaOH–2 mM EDTA and spin dialyzed. For larger plasmids (>15 kb), denatured DNA was neutralized with 0.2 M ammonium acetate (pH 4.8) and ethanol precipitated. Sequencing reactions were carried out with a U.S. Biochemical Sequenase sequencing kit on 5 μg of DNA, with [^{35}S]dATP from Amersham (Amersham, England). Universal primers (M13/pUC) were obtained from Boehringer (Mannheim, Germany), while further oligonucleotides were synthesized on a Beckman System 200 DNA synthesizer.

Reaction mixtures were separated on 6 to 8% acrylamide gels in 1× Tris borate buffer (as described by Sambrook et al. [30]). Gels were dried under vacuum and exposed to X-ray film at room temperature.

Sequence analysis. All sequences were analyzed by using the services of the Irish National Centre for Bioinformatics. Sequences from deletion points were entered into the Staden package (34) and aligned (by at least 20-bp overlap between each point) to give the consensus sequence. The sequence was analyzed with the PCGENE package (release 6.5). Similarity searches were performed at the National Center for Biotechnology Information (Bethesda, Md.), using the BLAST network (2). Multiple alignments were performed with the CLUSTAL program (18).

Nucleotide sequence accession number. The nucleotide sequence accession number of the sequence reported is Z34287.

RESULTS

Sequence analysis of the control region of PBSX. Earlier studies have shown that the PBSX repressor gene *xre* is located on restriction fragment x1, and our evidence suggests that the late promoter may be positioned on restriction fragment 38 (40, 41) (Fig. 1). A restriction map of the 5.2-kb region between the repressor gene *xre* and fragment 38 was generated (Fig. 1b). This region was found to contain a 1.0-kb *Eco*RI fragment, 316, which had been suggested to lie in the early region of PBSX (40). The DNA sequences of fragments x2, 316, x3, and 38 were determined and are shown together with the previously determined sequence of x1 (Fig. 2).

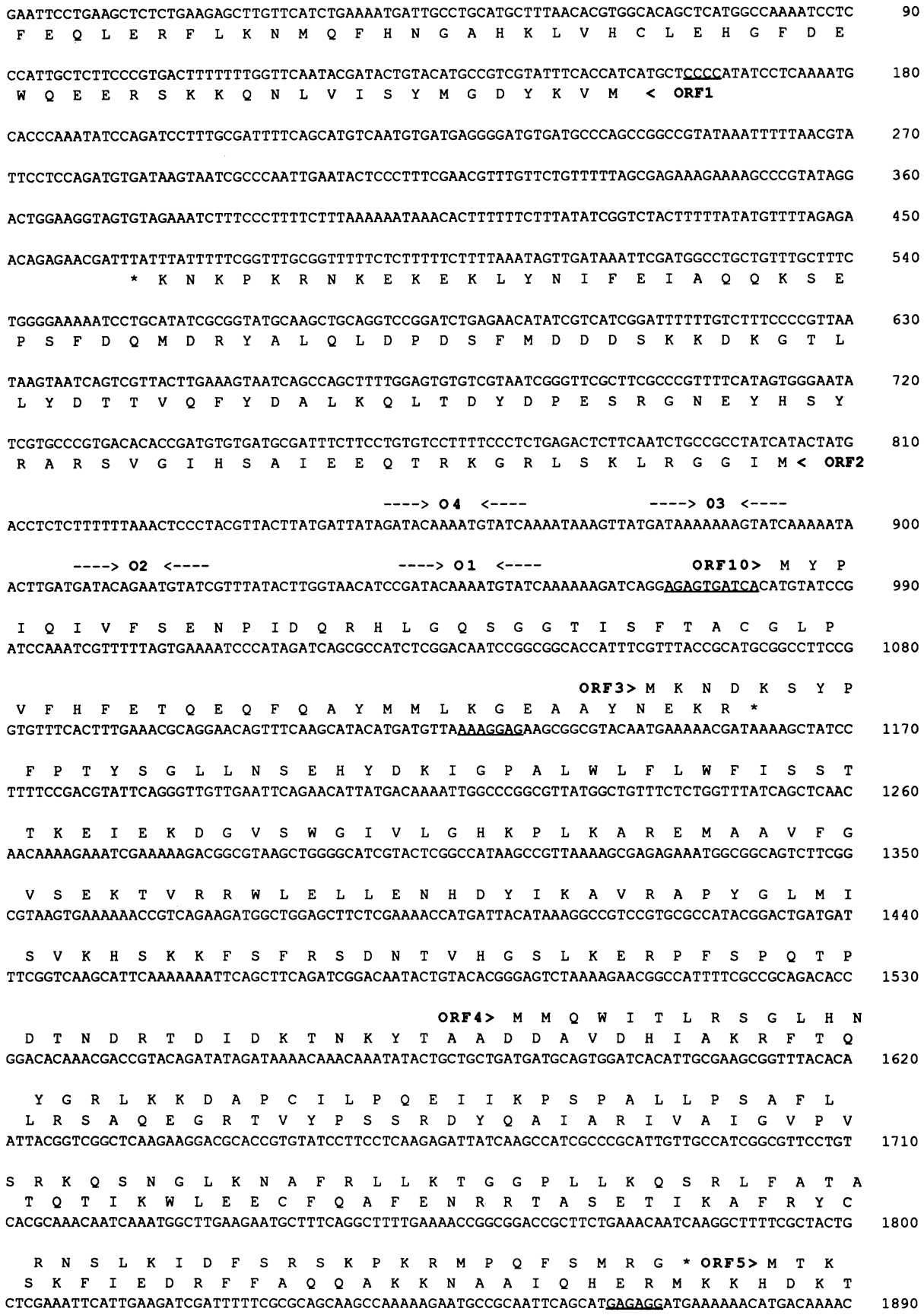
The 5,245-bp sequence was searched for ORFs of greater than 200 bp in length. Ten ORFs, ranging in size from 243 to 831 bp, were identified (Fig. 1 and 2). With the exception of ORF4, all ORFs had potential AUG or GUG start codons preceded by apparent Shine-Dalgarno sequences (16, 23). These ORFs were translated and each used to search the NBRF/PIR protein (release 37.0) and GenBank (release 78.0) translated databases. ORF1 and ORF9 are incomplete. ORF2 was previously identified as the PBSX repressor gene *xre* (41). ORF4, ORF6, and ORF10 showed no significant similarity to any known proteins in the databases searched.

ORF3 shows a putative HTH motif between residues 58 to 80, as determined by the method of Dodd and Egan (12), suggesting that ORF3 may be a DNA-binding protein.

ORF5 encodes a potential protein of 30.5 kDa and shows limited similarity to a number of NTP-binding proteins. Examination of the sequence revealed the presence of two short blocks of residues which are characteristic of NTP-binding proteins (11) (Fig. 3). Weak similarity was also found between the N-terminal region of ORF5 and the λ N protein. Closer analysis of the region led to the identification of a putative arginine-rich subdomain which is conserved in many RNA-binding proteins (Fig. 3) (20). This motif has been identified in the N-like gene products of the lambdoid phages and is essential for interaction of these antiterminators with their respective *nut* sites.

ORF7 showed similarity to 10 RNA polymerase sigma factors or putative sigma factors. These included σ^B, σ^F, and σ^K of *B. subtilis* (4, 13, 42), σ^{35} and σ^{28} of *Bacillus thuringiensis* (1), WhiG and RpoX of *Streptomyces coelicolor* (8, 19), RpoX

FIG. 2. Nucleotide sequence of 5,245 bp of the early and middle regions of PBSX. The deduced amino acid sequences of putative ORFs larger than 200 bp are shown above the DNA sequence. The first 1,200 bp was previously reported by Wood et al. (41). Proposed ribosome binding sites for these ORFs are underlined. Inverted repeats identified are indicated as arrows above the sequence, and direct repeats are indicated as double lines. The previously identified operator sites (O1 to O4) are indicated. T_1 indicates an inverted repeat structure which could form a stem-loop structure important for transcription termination.


GAATTCCTGAAGCTCTCTGAAGAGCTTGTTCATCTGAAAATGATTGCCTGCATGCTTTAACACGTGGCACAGCTCATGGCCAAAATCCTC 90
F E Q L E R F L K N M Q F H N G A H K L V H C L E H G F D E
CCATTGCTCTTCCCGTGACTTTTTTTGGTTCAATACGATACTGTACATGCCGTCGTATTTCACCATCATGCTCCCCATATCCTCAAAATG 180
W Q E E R S K K Q N L V I S Y M G D Y K V M < ORF1
CACCCAAATATCCAGATCCTTTGCGATTTTCAGCATGTCAATGTGATGAGGGGATGTGATGCCCAGCCGGCCGTATAAATTTTTAACGTA 270
TTCCTCCAGATGTGATAAGTAATCGCCCAATTGAATACTCCCTTTCGAACGTTTGTTCTGTTTTTAGCGAGAAAGAAAAGCCCGTATAGG 360
ACTGGAAGGTAGTGTAGAAATCTTTCCCTTTTCTTTAAAAAATAAACACTTTTTTCTTTATATCGGTCTACTTTTTATATGTTTTAGAGA 450
ACAGAGAACGATTTATTTATTTTTCGGTTTGCGGTTTTTCTCTTTTTCTTTTAAATAGTTGATAAATTCGATGGCCTGCTGTTTGCTTTC 540
* K N K P K R N K E K E K L Y N I F E I A Q Q K S E
TGGGGAAAAATCCTGCATATCGCGGTATGCAAGCTGCAGGTCCGGATCTGAGAACATATCGTCATCGGATTTTTTGTCTTTCCCCGTTAA 630
P S F D Q M D R Y A L Q L D P D S F M D D D S K K D K G T L
TAAGTAATCAGTCGTTACTTGAAAGTAATCAGCCAGCTTTTGGAGTGTGTCGTAATCGGGTTCGCTTCGCCCGTTTTCATAGTGGGAATA 720
L Y D T T V Q F Y D A L K Q L T D Y D P E S R G N E Y H S Y
TCGTGCCCGTGACACACCGATGTGTGATGCGATTTCTTCCTGTGTCCTTTTCCCTCTGAGACTCTTCAATCTGCCGCCTATCATACTATG 810
R A R S V G I H S A I E E Q T R K G R L S K L R G G I M < ORF2
----> O4 <---- ----> O3 <----
ACCTCTCTTTTTTAAACTCCCTACGTTACTTATGATTATAGATACAAAATGTATCAAAATAAAGTTATGATAAAAAAAGTATCAAAAATA 900
----> O2 <---- ----> O1 <---- ORF10> M Y P
ACTTGATGATACAGAATGTATCGTTTATACTTGGTAACATCCGATACAAAATGTATCAAAAAAGATCAGGAGAGTGATCACATGTATCCG 990
I Q I V F S E N P I D Q R H L G Q S G G T I S F T A C G L P
ATCCAAATCGTTTTTAGTGAAAATCCCATAGATCAGCGCCATCTCGGACAATCCGGCGGCACCATTTCGTTTACCGCATGCGGCCTTCCG 1080
ORF3> M K N D K S Y P
V F H F E T Q E Q F Q A Y M M L K G E A A Y N E K R *
GTGTTTCACTTTGAAACGCAGGAACAGTTTCAAGCATACATGATGTTAAAAGGAGAAGCGGCGTACAATGAAAAACGATAAAAGCTATCC 1170
F P T Y S G L L N S E H Y D K I G P A L W L F L W F I S S T
TTTTCCGACGTATTCAGGGTTGTTGAATTCAGAACATTATGACAAAATTGGCCCGGCGTTATGGCTGTTTCTCTGGTTTATCAGCTCAAC 1260
T K E I E K D G V S W G I V L G H K P L K A R E M A A V F G
AACAAAAGAAATCGAAAAAGACGGCGTAAGCTGGGGCATCGTACTCGGCCATAAGCCGTTAAAAGCGAGAGAAATGGCGGCAGTCTTCGG 1350
V S E K T V R R W L E L L E N H D Y I K A V R A P Y G L M I
CGTAAGTGAAAAAACCGTCAGAAGATGGCTGGAGCTTCTCGAAAACCATGATTACATAAAGGCCGTCCGTGCGCCATACGGACTGATGAT 1440
S V K H S K K F S F R S D N T V H G S L K E R P F S P Q T P
TTCGGTCAAGCATTCAAAAAAATTCAGCTTCAGATCGGACAATACTGTACACGGGAGTCTAAAAGAACGGCCATTTTCGCCGCAGACACC 1530
ORF4> M M Q W I T L R S G L H N
D T N D R T D I D K T N K Y T A A D D A V D H I A K R F T Q
GGACACAAACGACCGTACAGATATAGATAAAACAAACAAATATACTGCTGCTGATGATGCAGTGGATCACATTGCGAAGCGGTTTACACA 1620
Y G R L K K D A P C I L P Q E I I K P S P A L L P S A F L
L R S A Q E G R T V Y P S S R D Y Q A I A R I V A I G V P V
ATTACGGTCGGCTCAAGAAGGACGCACCGTGTATCCTTCCTCAAGAGATTATCAAGCCATCGCCCGCATTGTTGCCATCGGCGTTCCTGT 1710
S R K Q S N G L K N A F R L L K T G G P L L K Q S R L F A T A
T Q T I K W L E E C F Q A F E N R R T A S E T I K A F R Y C
CACGCAAACAATCAAATGGCTTGAAGAATGCTTTCAGGCTTTTGAAAACCGGCGGACCGCTTCTGAAACAATCAAGGCTTTTCGCTACTG 1800
R N S L K I D F S R S K P K R M P Q F S M R G * ORF5> M T K
S K F I E D R F F A Q Q A K K N A A I Q H E R M K K H D K T
CTCGAAATTCATTGAAGATCGATTTTTCGCGCAGCAAGCCAAAAAGAATGCCGCAATTCAGCATGAGAGGATGAAAAAACATGACAAAAC 1890

```
R T I E Q I L D E L R R G R R P L L A D K P A E S D A S R Y D
  N N R T D F G R A E K R E T S I T G G Q T G R I R R K Q V *
GAACAATCGAACAGATTTTGGACGAGCTGAGAAGAGGGAGACGTCCATTACTGGCGGACAAACCGGCCGAATCAGACGCAAGCAGGTATG   1980

  C L R C K D Q G G Y L V R Q N G L E V W T M C S C M A E R
ATTGCCTGCGCTGCAAGGATCAGGGAGGTTATCTCGTCAGGCAGAATGGCCTGGAAGTCTGGACGATGTGCAGCTGCATGGCAGAACGAA   2070

K V K R L L G A S E I T H A F R Q L G F K E F R T E G K P Q A
AAGTGAAGCGGCTGCTGGGTGCAAGTGAGATTACCCACGCTTTCAGACAGCTCGGCTTCAAGGAATTCCGCACGGAGGGAAAGCCGCAGG   2160

  I K D A F E C T K E Y V A D Y E Q I K D C R K N S I A L L
CCATAAAAGACGCATTTGAGTGTACAAAAGAGTATGTTGCCGATTATGAACAAATCAAGGATTGCCGAAAAAACAGCATTGCCCTTTTAG   2250

G Q P G S G K T H L L T A A A N E L M R T C Y V P V I Y F P F
GACAGCCCGGATCAGGGAAAACACACCTTTTGACCGCCGCAGCCAACGAATTAATGAGAACATGCTATGTGCCTGTCATTTATTTTCCGT   2340

  V E G F T D L K N D F A L L E A K L N R M K Q A D V L F I
TCGTGGAAGGCTTTACTGATCTGAAAAATGATTTTGCCCTATTAGAAGCGAAGCTGAACCGGATGAAGCAGGCGGATGTGCTGTTCATTG   2430

D D L F K P V N G K P R A T D W Q L E Q M Y S V L N Y R Y L N
ATGATCTGTTTAAACCGGTTAACGGCAAACCCCGCGCTACAGATTGGCAATTAGAGCAAATGTATTCGGTACTCAACTATCGCTACTTAA   2520

  H K P I L L S S E L T I E G L V R V D E A L G T R I Y E M
ATCATAAACCGATTTTGCTTTCGAGCGAGCTGACAATTGAAGGACTTGTACGGGTCGATGAAGCGCTCGGCACGAGAATCTATGAGATGT   2610

C S D Y L V I I K G A A Y E L N H R L E G V R *
GCAGTGACTATTTAGTGATTATCAAAGGAGCAGCCTACGAGCTTAACCATAGATTGGAGGGCGTCAGATAATGTGTAAGCTTTGTCAAAC   2700

AAAGAAAGTCATTGTGGAACATACCGGTATTGGAGTTGTTTTTCATCCATGTCCGAACTGCCGGTCCGCGACTGACTTAACGCCTGTCAT   2790

TCAAAAGCTGGAGCAAATGCTGACAGCGGGAAAAGCGAGGCTGAATATCTATGATTAAACAGCTGACTGCTCTAATCGCTTTGCTGTTTC   2880

                                        ORF6>  V K V L A K T K Q A E K S
GGGCAAAGCGAACAGAAAAAAACATTGAACAATGGTATAAGGACGACGGGAAGTGAAAGTGTTGGCAAAGACAAAACAGGCAGAGAAAAG   2970

  P A P W R A V P C G D T K P I Y I Y S A Y S E E E K E R F P
CCCTGCGCCGTGGCGTGCTGTCCCGTGCGGGGATACGAAACCGATCTATATTTATTCAGCTTACAGTGAAGAAGAAAAAGAAAGATTTCC   3060

  Y S N G R L I A A V F D L S S Y S Q K S N A S L M A A A P E
GTACTCAAACGGGCGGCTGATTGCAGCTGTATTTGACCTCAGCTCTTATTCGCAAAAAAGCAATGCCTCTTTGATGGCCGCTGCGCCTGA   3150

  L L E A S K A A V D F L K G N S I H S K E R I I Q L L E K A
ATTGCTGGAAGCGTCTAAAGCAGCAGTTGATTTTCTGAAAGGGAATTCTATTCATTCAAAGGAGCGTATCATTCAGCTATTAGAAAAAGC   3240

  E A S A A P K R G G N K T *                              =======          -
TGAAGCAAGCGCTGCACCGAAAAGGGGAGGAAATAAAACATGATTCATCCGAAAAAACTGCTGCATATCGATTCCGTCACGCTTAAGAGC   3330

--><---          =======                    ========              ========
CAGCTGGAGGACGGGAAAATCCGTCATTATTGTGGACGGCATCAAGCAAGAAGCATGGATCACAGAAGCGCCAGAGCATGGAAAAACGCT   3420

CGTCGAAACAAGAAAGGGCGATCTTGCTCGTGTGGAATTTGAAATCGGCTACAAATTAAATTAAAGCGAAAACAGAATACGTCCAAGACG   3510

        ---------------------->T1<----------------------                    ORF7>  M
GAAAGCCTGCGGACACTGATCAACTGCACAGCATTTGTGCGTTGATTGGTGTCCGTTTTTTATTTGCCAAAAATGAGGAGGATCATAGAA   3600

  Q D L L F E Y K R T L K Q T R I Q Y K P L A E A D E S V L
TGCAAGACTTACTATTTGAATATAAACGCACGCTCAAACAAACAAGAATACAATATAAACCGCTCGCTGAGGCAGATGAATCCGTGCTCT   3690

S A E E L K D K K I I R N M I T D L E Y V T E W L E K G R Q P
CAGCTGAAGAGCTGAAGGATAAAAAAATCATCAGAAATATGATTACTGATCTTGAATATGTAACAGAATGGCTTGAAAAAGGAAGGCAGC   3780

  G I R R A I D R R D V Y Q R L M I K D P R I I E S F S S A
CCGGCATCAGACGGGCGATTGACCGGCGTGATGTTTACCAGCGGCTGATGATCAAGGACCCGAGAATCATCGAATCATTTTCCAGCGCTA   3870
```

FIG. 2—*Continued.*

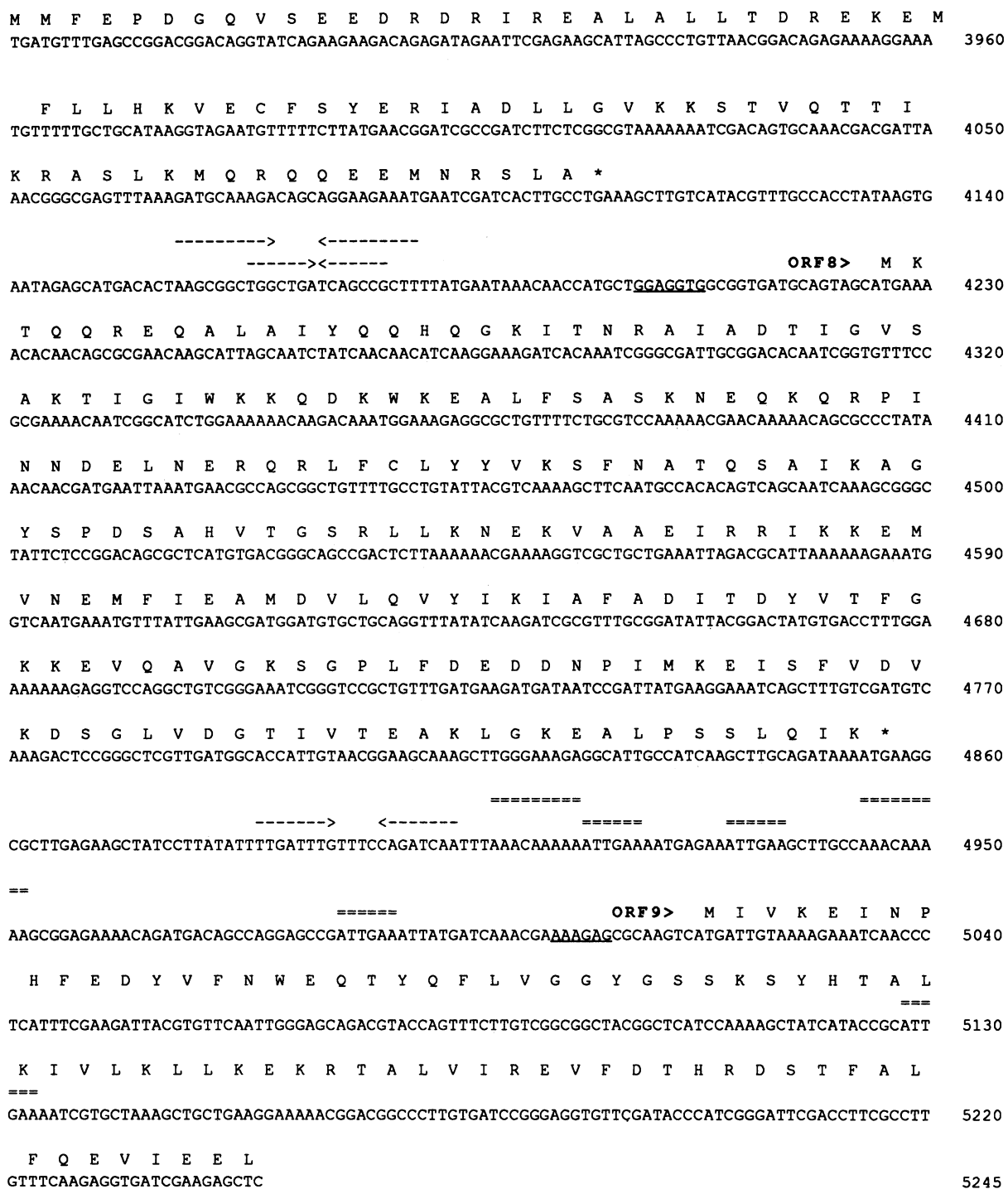


FIG. 2—*Continued.*

of *Streptomyces aureofaciens* (19), WhiG of *Streptoverticillium griseocarneum* (33), σ^F of *Pseudomonas aeruginosa* (35), and σ^G of *Clostridium acetobutylicum* (32).

Multiple alignments of known sigma factors have revealed regions and subregions of high conservation (15, 17, 22). Region 4.2 usually has an HTH DNA-binding motif which interacts with the −35 region of the promoter. A sequence which gave a score of 4.0 on the Dodd and Egan weight matrix test for HTH motifs was identified in ORF7 between residues 130 and 152. The likelihood of this region forming an HTH is high (12). This HTH motif is preceded by a short stretch of amino acids which are conserved across sigma factors in region 4.1 (Fig. 4).

ORF8 encodes a potential protein of 23.6 kDa and is 30% identical to the small subunit of terminase (gp1) from phage SPP1 of *B. subtilis* and to an ORF downstream of the *rinA* gene

	A		B
NTP binding motif	G/AXXXXGKT/S	40-100	XXXXXD (3 hydrophobic)
Coronavirus B3	GSPGAGKS	35	VVIMDD
ORF5	GQPGSGKT	48	DVLFI D

λN	qTRRRERRaEK
HIV Tat	RkKRRgRRRap
Bacillus L14	RTKRguRRpDg
E. coli S6	KDeRRERRDDf
Consensus	BOBRBJRRRZB
ORF5	dElRRgRRpll

FIG. 3. Putative motifs identified in ORF5. HIV, human immunodeficiency virus. Notation for the consensus sequence: B, basic, O, nonbasic; R, arginine; J, acidic; Z, charged amino acid.

in ϕ11 of *Staphylococcus aureus* (37, 43). A putative HTH motif was identified in ORF8 (residues 21 to 41), scoring 4.48 according to Dodd and Egan (12). The gp1 protein also has an HTH motif at the equivalent position. In addition, a putative ATP-binding motif (motif A, AXXXXGPL; residues 159 to 166) and a nucleotide-binding pocket (motif B, DE; residues 66 and 67) were identified in ORF8; these are also present in gp1. These motifs were also identified in the incomplete ORF downstream from the *rinA* gene in ϕ11 of *S. aureus* (43). We propose that ORF8 also encodes a terminase subunit. The gene organization of the terminase subunits is highly conserved in SPP1 and in several *E. coli* bacteriophages (5, 37). In these phage systems, the terminase enzyme, consisting of a small and a large subunit, recognizes a *pac* site which is located within the small subunit gene and is flanked by repeated elements. No putative *pac* site was identified in this region. ORF9, which is incomplete, is 43% identical to the amino-terminal region of the large subunit (gp2) of the terminase from SPP1.

Transcription from fragment 38 is under positive control. A Campbell-type integration mediated by fragment 38 is nonmutagenic; it was hypothesized that this fragment contains a promoter for the late genes (40). To test this hypothesis, fragment 38 (a 1.3-kb *Eco*RI-*Sac*I fragment) was cloned immediately upstream of a reporter gene, *lacZ*, in pDG268. This construct was integrated via a double-crossover event at the *amyE* locus on the chromosome of *B. subtilis* RB1081 (7), a strain which is deleted for the PBSX prophage, creating *B. subtilis* RB1081::38-*lacZ*. Hybridization analysis of *B. subtilis* RB1081 showed that the region deleted for PBSX includes all the sequences cloned by Wood et al. (40), including the *xre* gene and the genes of the late operon (data not shown). *B. subtilis* RB1081::38-*lacZ* did not show β-galactosidase activity, indicating either that fragment 38 does not have a promoter or that the promoter requires a PBSX function in *trans* (Table 2). To distinguish between these two possibilities, the same construct was integrated at the *amyE* locus in a PBSX$^+$ strain, *B. subtilis* IA420. In *B. subtilis* IA420::38-*lacZ*, expression of β-galactosidase was observed only after the addition of mitomycin, which induces PBSX. These results suggest that there is a promoter on fragment 38 which is regulated by a positive control factor, Pcf, synthesized after treatment with mitomycin and presumably encoded by PBSX.

TABLE 2. β-Galactosidase activities of integrant strains tested for positive control from fragment 38

B. subtilis strain	β-Galactosidase activity[a]	
	−Mitomycin	+Mitomycin
IA420::38-*lacZ* (wild type)	−	+
RB1081::38-*lacZ* (deletion)	−	−
RB1081::38-*lacZ*/pEB112	−	ND
RB1081::38-*lacZ*/pEB112-x4	+	ND

[a] Screening of activity on colonies growing on medium containing X-Gal. ND, not determined.

Identification of the *pcf* gene. An experiment was designed to isolate the gene encoding the positive control factor (Pcf). This was based on the proposition that the *pcf* gene should work in *trans* and confer β-galactosidase activity on the PBSX$^-$ strain *B. subtilis* RB1081::38-*lacZ*. pHV1435h contains 15 kb of DNA spanning the middle region of PBSX. This DNA was digested with *Hin*dIII, and fragments were ligated into the *Hin*dIII site downstream from the weak *tac* promoter in the *E. coli-B. subtilis* shuttle vector pEB112 (21). Ligation mixtures were introduced into *B. subtilis* RB1081::38-*lacZ*, selecting for kanamycin resistance and screening for β-galactosidase activ-

```
BS SigB   199 LQSVLHVLSDREKQIIDLTY----IQNKSQKETGDILGISQMHVSRLQRKAVKKL  248
BS SigF   197 LKEAISDLEEREKLIVYLRY----YKDQTQSEVAERLGISQVQVSRLEKKILKQI  245
BS SigK   149 VKQYIDILDDREKEVIVGRFGLDLKKEKTQREIAKELGISRSYVSRIEKRALMKM  202
BT Sig35  150 LMKALHQLNDREKQIMELRFGLAGGEEKTQKDVADMLGISQSYISRLEKRIIKRL  204
BT Sig28  149 IKEYIDILDEREKEVIVKRFGLGLDKEKTQREIAKALGISRSYVSRIEKRALMKM  203
SC WhiG   221 LARAINTLPEREKTVVTLYY----YEGLTLAEIGNVLGVTESRVSQIHTKSVLQL  270
SC RpoX   229 VRHLLVQLPEREQRILLLRY----YSNLTQSQISAELGVSQMHVSRLLARSFQRL  279
PA SigF   187 LADAIAKLPERERLVLALYY----DEELNLKEIGEVLGVSESRVSQLHSQCAARL  237
CA SigG   180 IKEAMKKLSDREKMILNMRF----FDGRTQMEVADEIGISQAQVSRLEKTALKHM  230
ORF7      108 -REALALLTDREKEM----FLLHKVECFSYERIADLLGVKKSTVQTTIKRASLKM  157
                  . * .**. .    .           .   .*.   .          .
                  ---------------  ----------------------------
                       4.1                  4.2
                                    -------   ------------
                                      H   T       H
```

FIG. 4. Multiple alignment of the deduced amino acid sequence for ORF7 (Pcf) and the sigma factors *B. subtilis* (BS) SigB, SigF, and SigK (SpollC), *B. thuringiensis* (BT) Sig35 and Sig28, *Streptomyces coelicolor* (SC) WhiG and RpoX, *P. aeruginosa* (PA) SigF, and *C. acetobutylicum* (CA) SigG. The alignment was produced by using CLUSTAL (18). Asterisks underlie residues conserved in all proteins, and dots underlie conserved residues. The position of conserved regions 4.1 and 4.2 (including the HTH motif) are indicated below the sequence, and residues in boldface have previously been identified in multiple alignments of sigma factors (22).

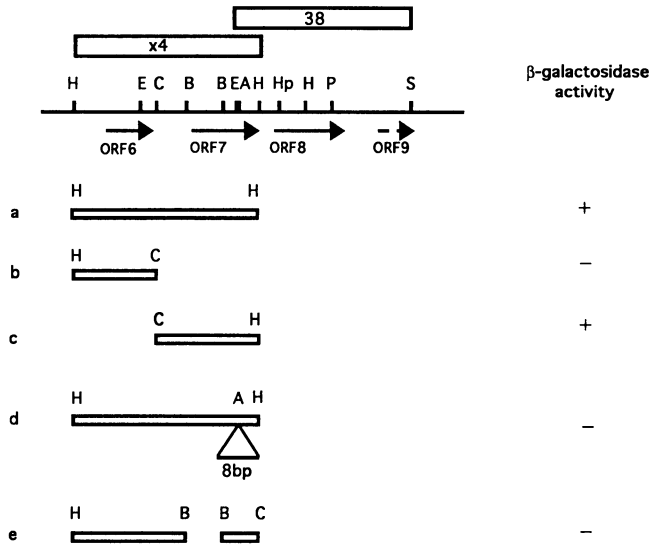


FIG. 5. Identification of a positive control factor (Pcf) required for transcription from the PBSX late promoter (P_L). The restriction map of the fragment x4/38 region of PBSX and the identified ORFs in this region are shown. The DNA fragments, a to e, were cloned into the *E. coli-B. subtilis* shuttle vector pEB112 (21) and transformed into *B. subtilis* RB1081::38-*lacZ*. Transformants were screened for β-galactosidase activity as described in Materials and Methods. β-galactosidase activity was observed in the presence of fragment x4 in *trans* (a), suggesting that either or both ORF6 and ORF7 were required for transcription from fragment 38. Fragments b and c were subsequently subcloned and when tested indicated that ORF7 was required only for positive control. An insertion of 8 bp into a unique *Hpa*I (HP) site in fragment x4 disrupts ORF7 (d). The deletion of a *Bcl*I (B) fragment, including the 5′ end of ORF7 (e) also disrupted positive control. Other restriction sites: H, *Hin*dIII; E, *Eco*RI; C, *Cla*I; B, *Bcl*I; A, *Hpa*I; P, *Pst*I; S, *Sac*I.

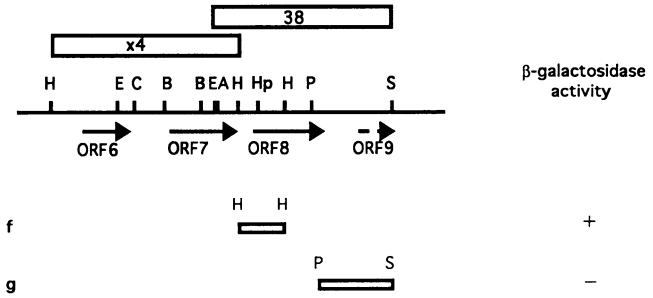


FIG. 6. Identification of promoter activity in fragment 38. The restriction map of the fragment x4/38 region of PBSX and the identified ORFs in this region are shown. To narrow down the late promoter (P_L) region in fragment 38, the two fragments shown (f and g) were subcloned into pDG268 and integrated via a double-crossover event into the α-amylase gene of *B. subtilis* RB1081. Integrant strains were tested for β-galactosidase activity in the presence of fragment x4 in *trans*. Results are shown in the right-hand column and indicated that P_L is within a 358-bp *Hin*dIII fragment (f). Restriction sites are as defined in the legend to Fig. 5.

ity. Three transformants, each of which grew poorly, were isolated, and the colonies were sectored. When the colonies were restreaked onto indicator plates (as described above) they segregated white colonies. Each of the three clones had recombinant plasmids containing a 1.5-kb *Hin*dIII fragment insert called x4. The result suggested that fragment x4 carries the gene *pcf* for the positive control factor. Southern analysis showed that fragment x4 overlaps both fragments 316 and 38, as indicated in Fig. 1 (results not shown).

Mutational analysis of x4. Fragment x4 includes ORF6 and ORF7. Cleavage of x4 with *Cla*I produces two fragments, one carrying ORF6 and the other carrying ORF7 (Fig. 5). Each was subcloned into the *Sma*I site of pEB112, downstream from the weak *tac* promoter and in the correct orientation for expression (results not shown). Subsequent plasmids were transformed into *B. subtilis* RB1081::38-*lacZ*. Plasmids containing ORF6 alone did not confer β-galactosidase activity, but those containing ORF7 did (Fig. 5b and c). The predicted sizes of proteins encoded by ORF6 and ORF7 are 12.6 and 20.1 kDa, respectively. Both of these fragments were cloned into *E. coli* expression vectors and were found to encode proteins of the predicted sizes (results not shown).

ORF7 was mutagenized in two different ways to verify that it encoded the Pcf activity. First, pUC19/x4 was linearized with *Hpa*I and religated in the presence of an 8-bp *Xba*I phosphorylated linker (CTCTAGAG; Pharmacia). This construct (Fig. 5d) causes a frameshift mutation in the 3′ end of ORF7, which should lead to premature termination 17 amino acids downstream from the insertion. The *Hin*dIII insert carrying the mutagenized ORF7 was subcloned into the *Hin*dIII site of pEB112 and used to transform *B. subtilis* RB1081::38-*lacZ*. It did not confer β-galactosidase activity.

A second mutation was made in ORF7. pUC19/x4 was digested with *Bcl*I to release a unique 300-bp fragment which overlaps the start of ORF7 (Fig. 4e). The vector was religated, and the *Hin*dIII insert was recloned into pEB112 and used to transform *B. subtilis* RB1081::38-*lacZ*. This ORF7 deletion mutant did not confer β-galactosidase activity. It is concluded that ORF7 encodes the PBSX positive control factor Pcf.

Localization of a late promoter in fragment 38. As shown above, fragment 38 contains a promoter which can be activated in the presence of Pcf. A number of constructs were prepared in order to map the late promoter P_L more precisely.

pDG268/38 was completely digested with *Eco*RI and partially digested with *Pst*I to remove the 5′ 800 bp of fragment 38. Staggered ends were removed with T4 DNA polymerase, and the vector was religated, leaving fragment g upstream of the promoterless *lacZ* gene (Fig. 6). pDG268/g was integrated via a double-crossover event as described before into *B. subtilis* RB1081. No β-galactosidase activity was observed in the presence or absence of the positive factor in *trans*.

Fragment f (a 360-bp *Hin*dIII fragment; Fig. 6) was subcloned into the *Hin*dIII site of pDG268, again upstream of the promoterless *lacZ* gene and integrated into *B. subtilis* RB1081. Promoter activity was observed from this fragment only when the positive factor was present (Fig. 6).

These results suggested that the promoter in fragment 38 is located within the 360-bp *Hin*dIII fragment f. This fragment is directly downstream from the *pcf* gene.

The promoter was located more precisely by using primer extension analysis. Two hundred-microliter aliquots of competent cells of *B. subtilis* RB1081::38.f-*lacZ* were transformed with pEB112 and pEB112/x4; 100 μl was plated onto 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) plates. A second aliquot was added to 30 ml of LB broth. After 20 h of growth at 37°C, blue colonies were evident on X-Gal plates only in the case of pEB112/x4 transformants. RNA was isolated from 30 ml of the parallel cultures. Aliquots of these cultures were also tested on indicator plates for activity; those transformed with pEB112/x4 showed β-galactosidase activity.

A 20-base oligonucleotide (CAATCGCCCGATTTGTGA

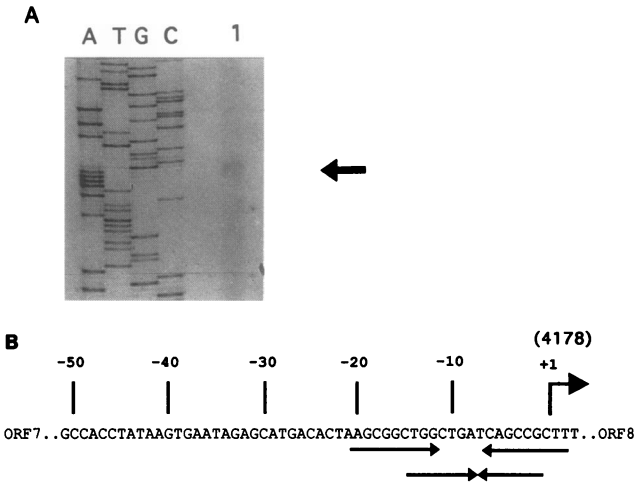


FIG. 7. (A) Mapping of the 5′ end of the PBSX late operon mRNA by primer extension analysis. The ^{32}P radioactively labelled oligonucleotide centered at position 4290 in Fig. 2 was hybridized to 30 μg of total RNA isolated from *B. subtilis* RB1081::38.f-*lacZ* transformed with pEB112/x4. Promoter activity had previously been localized to fragment f, as shown in Fig. 5. The primer extension reaction was analyzed on a 6% sequencing gel, in parallel to a sequencing reaction with the same oligonucleotide as the primer. Lanes A, T, G, and C indicate the sequencing reaction on fragment f. Lane 1 indicates the primer extension reaction. An arrow indicates the base representing the transcription start point, which is observed only in the presence of *pcf*. (B) Sequence of the complementary strand. The start point of transcription is indicated as +1, at position 4178 (see Fig. 2). Arrows below the sequence indicate inverted repeats previously identified.

TC) centered at position 4290 in Fig. 2 was used for primer extension on total RNA. Extensions were carried out at 50°C. The results are presented in Fig. 7. The apparent start site of transcription shown was mapped to position 4178 (with respect to Fig. 2) and was observed only in the presence of fragment x4 (controls not shown). This site lies between *pcf* and ORF8, 68 bp downstream of *pcf*.

DISCUSSION

A 5.2-kb region of PBSX DNA contains eight complete ORFs greater than 200 bp, including the previously identified repressor gene *xre* (41). One of these, ORF4, is contained within ORF3 and is not preceded by a ribosome binding site; it seems unlikely to encode a protein.

Six of the ORFs are in the same orientation. Only *xre* (ORF2), which is located at the extreme left-hand end of the region, is oriented in the opposite direction. Wood et al. (41) suggested that *xre* regulates the transcription of all other PBSX genes by binding to the putative operator sites which lie upstream from ORF10; they also proposed that *xre* is autoregulated and that there is divergent transcription from the region between *xre* and ORF10 from promoters recognized by the major vegetative form of RNA polymerase ($E^{\sigma A}$). The orientation of the other complete ORFs relative to *xre*, established by our sequence analysis, supports this proposal.

The data in the accompanying report (24) show that *xre* does indeed code for a protein which binds to the four operator sites between *xre* and ORF10 as suggested by Wood et al. (41). At low concentrations, it binds to two operator sites (O1 and O2) near ORF10, blocking both putative rightward promoters; at higher concentrations, it binds to O3 and O4 near *xre*, blocking both leftward putative promoters. These results strongly suggest that Xre represses rightward transcription of ORF10 at low concentrations and autoregulates transcription of *xre* at higher concentrations.

We suggest that ORF10, ORF3, ORF5, ORF6, and ORF7 are expressed as a single transcriptional unit, the middle genes. A Campbell-type insertion introduced by recombination in fragment 316 (Fig. 1) is polar, blocking all late gene expression (40). This implies that fragment 316 lies within a transcriptional unit and that a gene downstream from fragment 316 (possibly ORF7 or *pcf*) is required for late gene expression. Fragment 316 overlaps ORF5 and ORF6 (Fig. 1); thus, ORF5, ORF6, and ORF7 may belong to one transcriptional unit.

There is, however, a sequence (tentatively designated T_1) between ORF6 and ORF7 (*pcf*) which resembles a rho-independent transcriptional terminator (39). The putative terminator is 15 bp upstream of the ribosome binding site of ORF7 (*pcf*). If termination occurs at T_1, transcription of *pcf* may require the action of an antiterminator. The protein potentially encoded by ORF5 has some similarity to λ N protein and may thus encode such an antiterminator.

ORF7 encodes a positive control factor (Pcf) which is essential for transcription from P_L (Fig. 5). The sequence of Pcf shows similarity to sequences of several known sigma factors, including σ^B and σ^F of *B. subtilis* and WhiG *Streptomyces coelicolor*. The homology is strong in regions 4.1 and 4.2, which includes the segment of sigma factors which bind to the −35 region of each of the promoters (Fig. 4). It seems likely that Pcf directs the transcription of the late genes from P_L by acting as a sigma factor, but it remains puzzling that Pcf, at 20.1 kDa, is smaller than all known sigma factors.

Wood et al. (40) showed that integration directed by any fragment between fragment 38, at the left end, and fragment 39, which is 19 kb to the right, is mutagenic with respect to phibacin production. They proposed that the late genes were expressed from a single promoter in fragment 38. The data in Fig. 6 and 7 suggest that there is a promoter, P_L, within the left-hand end of 38 from which transcription starts at residue 4178. Integration directed by subfragment f is nonmutagenic (results not shown). Integration using the right-hand end subfragment g is mutagenic, implying that the end of an operon does not lie within subfragment g. The results suggest strongly that P_L is the only promoter for the late genes.

The first two late genes transcribed from P_L appear to be ORF8 and the incomplete ORF9. The predicted amino acid sequences of these ORFs show strong similarity to the sequences of the small and large subunits of the *B. subtilis* phage SPP1 terminase, respectively (37). Phage SPP1, P22, and P1 terminases initiate packaging from specific *pac* sites, thereafter cutting at sequences determined by a headful mechanism. PBSX does not package its own DNA but instead packages an apparently random array of bacterial DNA fragments of 13 kb (25, 26). It is possible that the PBSX terminase initiates packaging at an unidentified PBSX *pac* site in a manner similar to that of phages SPP1, P22, and P1 (5) and proceeds to package bacterial DNA in 13-kb fragments. If so, DNA sequences close to the PBSX locus should be packaged before those which are more distant. This prediction remains to be tested.

In summary, it appears that expression of the late genes of PBSX is regulated by a complex set of genes, including *xre* and *pcf*, which are sequentially expressed in a controlled manner in response to DNA damage. The repressor gene *xre* (24, 41) is absolutely required for the maintenance of the lysogenic state. If the Xre function in the temperature-sensitive mutant *xhi-1479* is eliminated at the nonpermissive temperature, or if cells

are exposed to DNA-damaging agents like mitomycin, PBSX is induced (6, 41) in a process which is dependent on the function of *recA* (27). We suggest that after elimination of Xre, transcription proceeds from two overlapping putative rightward promoters into ORF10. Expression of ORF5 may be required if transcription is to proceed into *pcf*. Expression of *pcf*, which may encode a sigma factor, is required for transcription from P_L and thus late gene expression.

Expression of PBSX is suicidal for the host and for PBSX; the host is lysed, and particles which cannot propagate PBSX are released. The particles kill *Bacillus* strains which do not carry PBSX. Closely related strains which are nonlysogenic for PBSX, but which are competing for resources, are vulnerable to killing by PBSX, thereby increasing the chances of survival of PBSX lysogens in the same environment. It is interesting to ask why are there so many genes involved in triggering transcription from P_L. It seems possible that this complex system has evolved under selection for failsafe action. A series of switches, interlocked through feedback mechanisms and finely tuned to various physiological signals, would ensure that expression of PBSX late genes, which leads inevitably to lysis, would be initiated only as a last and suicidal response to DNA damage. Those cells that commit suicide, however, produce a factor that kills closely related but nonidentical cells while not affecting siblings of PBSX producers that have not sustained lethal damage.

ACKNOWLEDGMENTS

We thank S. Krogh for many helpful discussions.
This work was supported by Novo Nordisk, Bagsvaerd, Denmark.

REFERENCES

1. **Adams, L. F., K. L. Brown, and H. R. Whiteley.** 1991. Molecular cloning and characterization of two genes encoding factors that direct transcription from a *Bacillus thuringiensis* crystal gene promoter. J. Bacteriol. **173**:3846–3854.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.
3. **Antoniewski, C., B. Savelli, and P. Stragier.** 1990. The *spoIIJ* gene which regulates early development steps in *Bacillus subtilis* belongs to the class of environmentally responsive genes. J. Bacteriol. **172**:86–93.
4. **Binnie, C., M. Lampe, and R. Losick.** 1986. Gene encoding the σ^{37} species of RNA polymerase σ factor from *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **83**:5943–5947.
5. **Black, L. W.** 1989. DNA packaging in dsDNA bacteriophages. Annu. Rev. Microbiol. **43**:267–292.
6. **Buxton, R. S.** 1976. Prophage mutation causing heat inducibility of defective *Bacillus subtilis* bacteriophage PBSX. J. Virol. **20**:22–28.
7. **Buxton, R. S.** 1980. Selection of *B. subtilis* 168 mutants with deletions of the PBSX prophage. J. Gen. Virol. **46**:427–437.
8. **Chater, K. F., C. J. Bruton, K. A. Blaskitt, M. J. Buttner, C. Mendez, and J. D. Helman.** 1989. The developmental fate of *Streptomyces coelicolor* hyphae depends upon a gene product homologous with the motility sigma factor of *Bacillus subtilis*. Cell **59**:133–143.
9. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1972. Characterisation of plasmid transformation in *B. subtilis*: kinetic properties and the effect of DNA conformation. Mol. Gen. Genet. **167**:251–258.
10. **Contente, S., and D. Dubnau.** 1979. Non-chromosomal antibiotic resistance in bacteria: genetic transformation of *E. coli* by R factor DNA. Proc. Natl. Acad. Sci. USA **69**:2110–2114.
11. **Denver, T. E., M. J. Glynias, and W. C. Merrick.** 1987. GTP-binding domains: three consensus sequence elements with distinct spacing. Proc. Natl. Acad. Sci. USA **84**:1814–1818.
12. **Dodd, I. B., and J. B. Egan.** 1990. Improved detection of helix-turn-helix DNA binding motifs in protein sequences. Nucleic Acids Res. **13**:5019–5026.
13. **Errington, J.** 1987. Two separate functional domains in the sigma-subunit of RNA polymerase in *Bacillus subtilis*. FEBS Lett. **224**:257–260.
14. **Geiduschek, E. P.** 1991. Regulation of expression of the late genes of bacteriophage T4. Annu. Rev. Genet. **25**:437–460.
15. **Gribskov, M., and R. R. Burgess.** 1986. Sigma factors from *E. coli*, *B. subtilis*, phage SPO1, and phage T4 are homologous proteins. Nucleic Acids Res. **14**:6745–6763.
16. **Hagar, P. W., and J. C. Rabinowitz.** 1985. Translational specificity in *Bacillus subtilis*, p. 1–32. *In* D. Dubnau (ed.), The molecular biology of the bacilli, vol. 2. Academic Press, Inc., New York.
17. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. Annu. Rev. Biochem. **42**:839–872.
18. **Higgins, D. G., and P. M. Sharp.** 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene **73**:237–244.
19. **Kormanec, J., and L. Potuckova.** Unpublished data.
20. **Lazinski, D., E. Grzadzielska, and A. Das.** 1989. Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. Cell **59**:207–218.
21. **Leonhardt, H., and J. C. Alonso.** 1988. Construction of a shuttle vector for inducible gene expression in *E. coli* and *B. subtilis*. J. Gen. Microbiol. **134**:605–609.
22. **Lonetto, M., M. Gribskov, and C. A. Gross.** 1992. The σ^{70} family: sequence conservation and evolutionary relationships. J. Bacteriol. **174**:3843–3849.
23. **McConnell, D. J., B. A. Cantwell, K. M. Devine, A. J. Forage, B. M. Laoide, C. O'Kane, J. F. Ollington, and P. M. Sharp.** 1986. Genetic engineering of extracellular enzyme systems of bacilli. Ann. N.Y. Acad. Sci. **469**:1–17.
24. **McDonnell, G. E., and D. J. McConnell.** 1993. Overproduction, isolation, and DNA-binding characteristics of Xre, the repressor protein from the *Bacillus subtilis* defective prophage PBSX. J. Bacteriol. **176**:5831–5834.
25. **Okamoto, K., J. A. Mudd, J. Mangan, W. M. Huang, T. V. Subbaiah, and J. Marmur.** 1968. Properties of the defective phage of *Bacillus subtilis*. J. Mol. Biol. **34**:413–428.
26. **Okamoto, K., J. A. Mudd, and J. Marmur.** 1968. Conversion of *B. subtilis* DNA to phage DNA following mitomycin C induction. J. Mol. Biol. **34**:429–437.
27. **O'Kane, C., M. A. Stephens, and D. J. McConnell.** 1986. Integrable α-amylase plasmid for generating random transcriptional fusions in *Bacillus subtilis*. J. Bacteriol. **168**:973–981.
28. **O'Reilly, M., K. Woodson, B. C. A. Dowds, and K. M. Devine.** 1994. The citrulline biosynthetic operon *argC-F*, and a ribose transport operon, *rbs*, from *Bacillus subtilis* are negatively regulated by SpoOA. Mol. Microbiol. **11**:87–98.
29. **Rodriguez, R. L., and R. C. Tait.** 1983. Recombinant DNA techniques: an introduction. Addison Wesley Publishers Ltd., London.
30. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
32. **Sauer, U., A. Treuner, M. Buchholz, and F. Duerre.** Unpublished data.
33. **Soliver, J., C. Granozzi, K. A. Plaskitt, and K. E. Chater.** Unpublished data.
34. **Staden, R.** 1982. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucleic Acids Res. **10**:4731–4761.
35. **Starnbach, M. H., and S. Lory.** 1992. The *fliA* (*rpoF*) gene of *Pseudomonas aeruginosa* encodes an alternative sigma factor required for flagellin synthesis. Mol. Microbiol. **6**:459–469.
36. **Sussman, R., and F. Jacob.** 1962. Sur une systeme de represion thermosensible chez le bacteriophage d'*Escherichia coli*. C. R. Acad. Sci. **254**:1517.
37. **Tavares, P., M. A. Santos, R. Lurz, G. Morelli, H. deLencastre,**

and T. A. Trautner. 1992. Identification of a gene in *Bacillus subtilis* bacteriophage SPP1 determining the amount of packaged DNA. J. Mol. Biol. **225:**81–92.

38. **Thurm, P., and A. J. Garro.** 1975. Isolation and characterization of prophage mutants of the defective *Bacillus subtilis* bacteriophage PBSX. J. Virol. **16:**184–191.
39. **Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. **246:**40–41.
40. **Wood, H. E., M. T. Dawson, K. M. Devine, and D. J. McConnell.** 1990. Characterization of PBSX, a defective prophage of *Bacillus subtilis*. J. Bacteriol. **172:**2667–2674.
41. **Wood, H. E., K. M. Devine, and D. J. McConnell.** 1990. Characterisation of a repressor gene (xre) and a temperature-sensitive allele from the *B. subtilis* prophage, PBSX. Gene **96:**83–88.
42. **Wu, J.-J., R. Schuch, and P. J. Piggot.** 1992. Characterization of a *Bacillus subtilis* operon that includes genes for an RNA polymerase σ factor and for a putative DD-carboxpeptidase. J. Bacteriol. **174:**4885–4892.
43. **Ye, Z.-H., and C. Y. Lee.** 1993. Cloning, sequencing, and genetic characterization of regulatory genes, *rinA* and *rinB*, required for the activation of staphylococcal phage ϕ11 *int* expression. J. Bacteriol. **175:**1095–1102.