

Nucleotide sequence of the *spo0B* gene of *Bacillus subtilis* and regulation of its expression

(sporulation/promoter mapping/gene fusion)

JEAN BOUVIER*, PATRICK STRAGIER*, CÉLINE BONAMY†, AND JEKISIEL SZULMAJSTER†

*Institut de Microbiologie, Bât. 409, Université Paris-Sud, 91405 Orsay Cedex, France; and †Laboratoire d'Enzymologie du Centre National de la Recherche Scientifique, 91190 Gif-sur-Yvette, France

Communicated by Jesse C. Rabinowitz, July 30, 1984

ABSTRACT The *spo0B* gene is one of the genes involved in initiation of sporulation of *Bacillus subtilis*. This gene, previously cloned into the pHV33 shuttle vector, is expressed in *Escherichia coli* and *B. subtilis*. We have determined the sequence of 1118 base pairs (bp) of the DNA insert carrying the *spo0B* gene. The promoter sequence of this gene shows the canonical T-A-T-A-A-T region at 10 bp from the transcriptional start (-10 region) but an unusual sequence, T-T-T-C-T, in the -35 region. The nucleotide sequence shows an open reading frame encoding a 192-amino-acid polypeptide of *M*_r 22,542, which is close to the molecular weight of the *spo0B* product synthesized in *E. coli* minicells. To investigate the regulation of the *spo0B* gene under a variety of physiological conditions, we constructed an in-frame fusion between the *spo0B* promoter proximal region and the *lacZ* gene of *E. coli*. This hybrid gene was subsequently integrated into the *B. subtilis* chromosome, and the β -galactosidase activity was measured. It was found that the *spo0B* gene is preferentially expressed during exponential growth; it is not induced by exhaustion of the growth medium nor repressed by glucose.

The morphological and biochemical events of sporulation in *Bacillus subtilis* are under the control of a large number of genes scattered along the chromosome and are grouped into five classes (*spo0*, *spoII*, *spoIII*, *spoIV*, and *spoV*) according to the stage of sporulation they are affecting (1). Any mutation in the *spo0* genes blocks the overall sporulation process. It is now well established that this class of genes controlling the initiation of sporulation lies into nine loci designated *spo0A*, *spo0B*, etc. (2).

A number of methods recently have been developed for cloning spore genes and applied by different investigators for the isolation of several *spo0* genes (3-7). We have reported the cloning and expression in *Escherichia coli* of the *spo0B* gene of *B. subtilis* and have shown that this gene is carried on a DNA fragment of 1.1 kilobases (kb) (8). We have isolated several hybrid plasmids carrying this gene either in a single copy in both orientations or in a tandemly arranged double copy. All of these hybrid plasmids were able to transform the Spo⁻ Rec⁻ Bs72 strain to Spo⁺, which indicated that the *spo0B* gene in these hybrid plasmids is being transcribed by the *B. subtilis* RNA polymerase by using an internal promoter on the cloned DNA fragment (8).

For more information on the specific role of the *spo0B* gene in the initiation process of sporulation, it was of great interest to analyze the nucleotide sequence of this gene and the nature of its promoter and to compare these data to those obtained in other laboratories on two different zero-stage sporulation genes, *spo0F* from *B. subtilis* (9) and *spo0H* from *Bacillus licheniformis* (10). It also was of interest to study the expression of the *spo0B* gene and its regulation under a

variety of growth conditions. The results of these studies are reported here.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Media. *E. coli* MC4100 (11), grown in L broth (12), was used as a recipient strain for recombinant plasmids. *B. subtilis* BsB20 (*trpC2 phe-1*; GSY1070 from C. Anagnostopoulos) and its *spo0B* derivative (Bs31; JH648 from J. Hoch) were used for chromosomal integration of plasmid pOB110 and grown in DS medium (13). Appropriate antibiotics were added to the following final concentrations: ampicillin, 50 μ g/ml, and chloramphenicol, 3 μ g/ml.

DNA Manipulations. Plasmid purification and DNA-fragment isolation, digestion with restriction enzymes, ligation, and transformation were as described (8, 14). DNA sequence determination was carried out as described by Maxam and Gilbert (15). Products of the chemical degradation reactions were analyzed on the ultrathin denaturing gels of Sanger and Coulson (16). Transcriptional start was determined with S1 nuclease (Boehringer Mannheim) or reverse transcriptase (a gift from M. Yaniv) as described by Débarbouillé and Raibaud (17). *B. subtilis* mRNA was extracted and purified essentially as described by Gilman and Chamberlin (18).

Construction of a *spo0B-lacZ* Fusion. pSKS107 plasmid (19) was first deleted after partial *Ava* I digestion, which eliminated the distal part of the *lac* operon, known to be toxic in *B. subtilis* cells (20). A 915-base-pair (bp) *Hind*III fragment was purified from pGsOB10 plasmid (8). This fragment carried the *Hind*III-*Bam*HI part of pBR322 and the first 590 bp of the sequence given in Fig. 3. Insertion of this fragment in the unique *Hind*III site of pSKS107 Δ *Ava* I led to an in-frame fusion between *spo0B* and *lacZ* coding sequences (plasmid pOB108). A 1135-bp *Nae* I-*Taq* I fragment carrying a chloramphenicol-resistance (*Cm*^r) determinant was purified from plasmid pC194 (ordinate 975-2110 in ref. 21) and inserted between the *Sma* I and *Sal* I sites of pOB108. This plasmid, which can be amplified in *E. coli* but is unable to replicate in *B. subtilis* cells, was designated pOB110. Extensive restriction mapping confirmed that this plasmid had the expected structure, which is shown in Fig. 4.

β -Galactosidase Assays. β -Galactosidase assays were as described by Zuber and Losick (22), and the specific enzymatic activity was expressed in Miller units (12).

RESULTS

Nucleotide Sequence Analysis. Both strands of the cloned 1118-bp DNA fragment of *B. subtilis* carried by the pGsOB10 plasmid (8) were sequenced completely by the strategy outlined in Fig. 1. The results are given in Fig. 2. Previous results have indicated that the *spo0B* gene stretches

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: bp, base pair(s); kb, kilobase(s); *Cm*^r, chloramphenicol resistance.

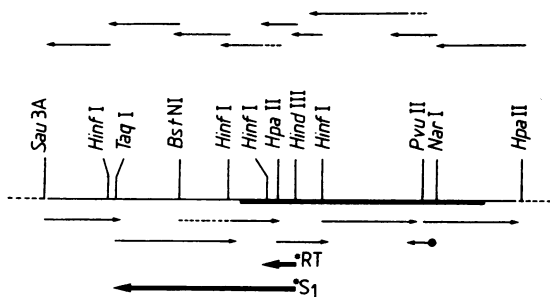


FIG. 1. Sequencing strategy of the *spo0B* gene. The arrows indicate the sites 5'-end-labeled with polynucleotide kinase, as well as the direction and extent of the sequences. The arrow with a circle denotes sequence analysis of a fragment labeled at its 3' end by using reverse transcriptase. Only the restriction sites used for labeling are shown. Probes used in the 5' mapping experiments (reverse transcriptase and S1 nuclease) are indicated by an asterisk. The heavy line represents the coding part of the *spo0B* gene.

on both sides of the unique *Pst* I site at the position 567 (7) and runs over a *Sau*3A site located between *Hind*III at 590 and the *Pvu* II site at 894 (8). There is only one open reading frame in the DNA insert that corresponds to these criteria and encodes the *spo0B* product. There are two ATG codons

for potential initiation of translation, one at position 444 and the other at position 462. The latter one, which is preceded at 7 bp by the G-G-A-G sequence complementary to the 3' terminus of the 16S rRNA (23, 24), is most likely the effective initiation translation codon of the *spo0B* gene. According to the rules of Tinoco *et al.* (25), the free energy of formation of this initiation site is -11.6 kcal/mol. The reading frame stops at position 1038 by an amber codon.

The amino acid sequence deduced from the nucleotide sequence is indicated on Fig. 2. This sequence corresponds to a 192-amino-acid polypeptide with a M_r of 22,542, which is in good agreement with that found previously in *E. coli* or *B. subtilis* minicell systems (8). This polypeptide is rich in charged residues (33.8%), with a light excess of acidic residues (34 versus 31 basic ones). Computer analysis has shown no significant amino acid homology with the products encoded by the two other sequenced *spo0* genes, *spo0F* from *B. subtilis* (9) and *spo0H* from *B. licheniformis* (10). There is also no homology to any other known proteins.

Transcriptional Start Site of the *spo0B* Gene. The promoter region of the *spo0B* gene was determined by mapping its transcriptional start site using mRNA purified from *B. subtilis* vegetative cells carrying the pGsOB10 plasmid (8). This mRNA was hybridized with either the *Hind*III-*Hinf*I 67-bp fragment or the *Hind*III-*Taq* I 432-bp fragment labeled at



FIG. 2. Nucleotide sequence of the *spo0B* gene and adjacent regions. Only the nontranscribed DNA strand is shown. Numbering is from the 5' end. The deduced amino acid sequence for the correct reading frame is given below. The putative ribosome binding site is underlined twice. Convergent arrows and dots indicate inverted repeats and their respective center of symmetry. The thick arrow shows the *spo0B* mRNA start; the corresponding "-10" and "-35" promoter signals are underlined with a thick line. The *Hind*III site used in the course of the construction of a hybrid *spo0B-lacZ* gene is shown.

their *Hind*III 5' terminus (see Fig. 1). The heteroduplex mixtures were treated, respectively, with reverse transcriptase or S1 nuclease. Fig. 3 shows the results of these mapping experiments along with the fragments generated by the chemical sequence analysis performed on the *Hind*III–*Taq*I probe. Comparison of the results obtained with both methods allowed unambiguous localization of the *spo*0*B* mRNA start at adenine-410. This transcriptional start is located 7 bp downstream of a T-A-T-A-A-T sequence, which matches the –10 consensus region recognized by the *B. subtilis* major vegetative RNA polymerase containing the σ^{55} subunit (27). This region is separated by 17 bp from a T-T-T-T-C-T sequence, which is different from the consensus –35 promoter region (T-T-G-A-C-A) found in the other *B. subtilis* genes sequenced so far (27). Upstream of the promoter region, there are several inverted repeat sequences that may control the expression of the *spo*0*B* gene (Fig. 2). Further upstream between nucleotides 148 and 181 there is a region preceding a run of thymidine residues that might favor the formation of stable loop structures ($\Delta G = -18$ kcal/mole) in the complementary RNA. This structure strongly resembles the rho-independent transcription stops described in *E. coli* (28), which might represent the termination signal of an adjacent gene.

Expression of a *spo*0*B*-*lacZ* Fusion. To study the regulation of the expression of the *spo*0*B* gene, we constructed the plasmid pOB110 in which the expression of the *E. coli lacZ* gene was placed under the control of transcriptional and translational signals of the *spo*0*B* gene. The details of the construction of this plasmid are described in *Materials and Methods* and in the legend of Fig. 4. This plasmid carries a 595-bp fragment from *spo*0*B* that allows the synthesis of a hybrid protein of which the first 45 amino acids are encoded by the

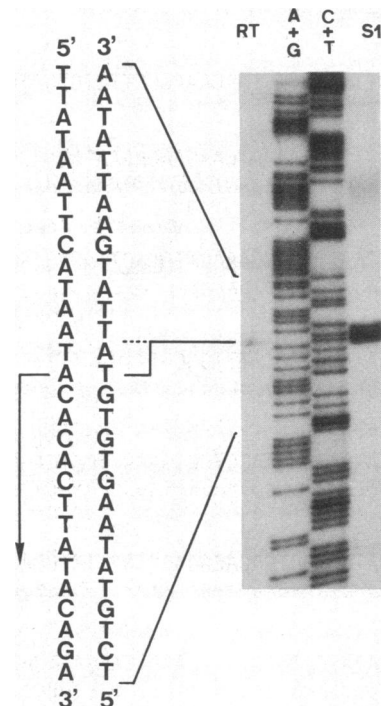


FIG. 3. Mapping of the 5' end of the *spo*0*B* transcript. After hybridization with vegetative mRNA from strain Bs31/pGsOB10 and treatment with reverse transcriptase (RT) or S1 Nuclease (S1), the DNA probes indicated in Fig. 1 were analyzed on a 7% polyacrylamide gel in the presence of 8 M urea. Part of the initial large probe was submitted to chemical reactions specific for purines (A+G) or pyrimidines (C+T) according to Maxam and Gilbert procedures (15). A 1.5-nucleotide correction has been made between the sequence ladder and the products of enzymatic reactions (26).

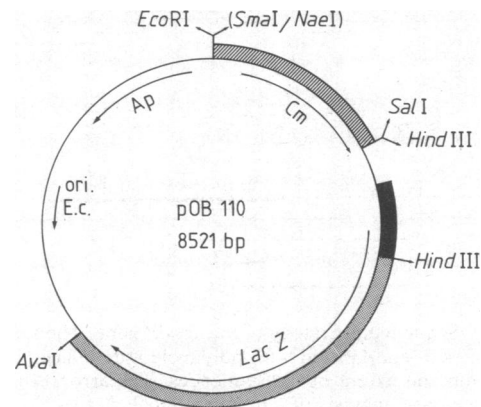


FIG. 4. Physical map of plasmid pOB110 carrying a *spo*0*B*-*lacZ* fusion. Arrows show the direction of transcription of the *Cm*^r and ampicillin (Ap)-resistance genes and of the hybrid *spo*0*B*-*lacZ* gene as well as the direction of replication in *E. coli* (*oriE.c.*). Thin lines indicate pBR322 sequences, the dotted region is from the *E. coli lac* operon, the striped region is from pC194 plasmid, and the black area indicates *spo*0*B* DNA. The *Sma*I and *Nae*I sites are not regenerated.

*spo*0*B* gene and which is endowed with β -galactosidase activity. This plasmid carries also the *Cm*^r marker, which can be expressed in *B. subtilis*. In the absence of a replication origin functional in *B. subtilis*, and since this plasmid carries only one region homologous with the chromosome, the selection for *Cm*^r cells requires a recombination event by a Campbell-like mechanism. The whole plasmid was integrated into the chromosome in a tandemly arranged fashion with the *spo*0*B* gene as detailed in ref. 22. This was first verified by cotransformation of the *Cm*^r marker with the *phe* marker

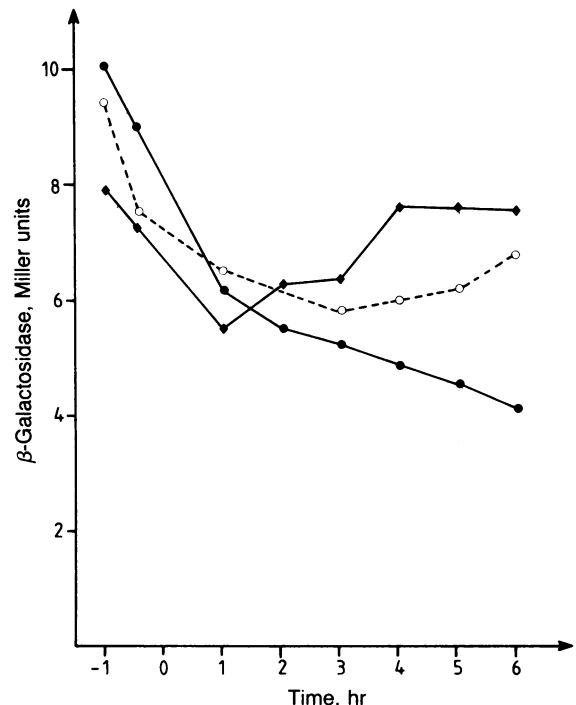


FIG. 5. β -Galactosidase activity in the chromosomal *spo*0*B*-*lacZ* fusions. The specific activity of β -galactosidase in Miller units (12) was measured at regular intervals during growth and stationary phase for the following strains containing chromosomally inserted pOB110: BsB20 grown in DS medium (\bullet) and in DS medium with 0.5% glucose (\circ) and Bs31 (*spo*0*B*) grown in DS medium (\blacklozenge). Results given are the average of three experiments.

closely linked to the *spo0B* gene on the *B. subtilis* chromosome (data not shown). Moreover, when using the asporogenic strain Bs31 carrying the OB136 mutation localized in the proximal part of the *spo0B* gene (7), both Spo⁺ and Spo⁻ Cm^r colonies were recovered, confirming integration of the pOB110 plasmid into the *spo0B* region. In that case, Spo⁻ Cm^r Lac⁺ clones (appearing as blue colonies on a medium containing 5-bromo-4-chloro-3-indolyl- β -D-galactoside) were purified and kept for further studies.

The β -galactosidase activity was measured under various conditions, and the results are summarized in Fig. 5. It appears that the expression of the *spo0B* gene was at maximum level during exponential growth and progressively decreased during sporulation until it reached about 40% of the value observed during vegetative growth. The addition of glucose to the growth medium did not affect the rate of synthesis of β -galactosidase. In the asporogenic strain, there was only a transitory decrease in the β -galactosidase activity at the beginning of the stationary phase, which progressively restored the level observed during vegetative growth.

DISCUSSION

The cloning of the *B. subtilis spo0B* gene has been the subject of studies in several laboratories (6–8). We have previously reported the cloning of this gene carried on a 1.1-kb fragment into the pHV33 shuttle vector. By using the minicell system of *E. coli* or *B. subtilis*, we have been able to show that the 1.1-kb DNA insert encodes a protein of a molecular weight of 24 kDa (8). In this paper we present the complete nucleotide sequence of the *spo0B* gene and the characterization of its transcriptional signal (Fig. 2). This sequence is in agreement with the genetic and restriction map described by Ferrari *et al.* (7) with the minor exception that the *Bal* I site is closer to the unique *Pst* I site than described by these authors (at position 536 in Fig. 2). The polypeptide deduced from the nucleotide sequence corresponds to 192 amino acids, and its calculated M_r of 22,542 is very close to that obtained in *E. coli* minicells. Our result is inconsistent with the 39-kDa polypeptide reported by Hirochika *et al.* (6) and assumed to be encoded by an *Eco*RI fragment carrying the *spo0B* gene cloned into the transducing phage ρ 11. The size of the *spo0B* gene product estimated from our current study is similar to those encoded by the *spo0F* gene of *B. subtilis* [173 amino acids (9)] or by the *spo0H* gene of *B. licheniformis* [168 amino acids (10)].

The knowledge of the primary structure of the sequenced *spo0* genes provides little information on their function and on the role they play in the initiation process of sporulation. It is generally assumed that the stage zero mutations have gross pleiotropic effects on the expression of genes involved in later stages of sporulation. This suggested that the products of the *spo0* genes are required for the progression of sporulation events (29, 30). Concerning specifically the *spo0A* and *spo0B* genes, it was reported by Trowsdale *et al.* (31) that at least some of their pleiotropic effects can be suppressed by mutations in the *abrB* locus, resulting in ribosomal alterations. Thus, it was suggested that the *spo0A* and *spo0B* gene products may interact with the ribosomes by favoring, for instance, the translation of mRNAs specific for sporulation. However, no significant homology has so far been found between the *spo0B* gene product and the known proteins involved in translation.

Numerous promoter regions have been examined in *B. subtilis* (27). In this organism the temporal changes in transcription associated with the sporulation process require a higher degree of complexity than is observed in *E. coli*. The compilation of data on promoter sequences together with the results obtained with *in vitro* transcription have revealed the existence of several classes of promoters recognized by different σ factors associated with the *B. subtilis* RNA poly-

merase (32). The promoter recognized by the major form of RNA polymerase containing σ^{55} is similar to that found in *E. coli*. The promoter found in the *spo0B* gene is apparently of the same nature; it contains the canonical sequence T-A-T-A-A-T in its -10 region upstream from the transcription start site. However, the -35 region seems to be less typical and shows a T-T-T-T-C-T sequence located at 17 bp from the -10 region or a T-T-G-T-T-T sequence at a distance of 21 bp, which is unusual for *B. subtilis*. In fact, by applying the rules established for the promoters of *E. coli* (33), it appears that these two sequences are more or less equivalent with respect to their homology score values (69.2 for the first sequence and 68.0 for the second). Compared to the *E. coli* promoters, these values are relatively high. No promoter sequences recognized by the minor forms of σ factors designated σ^{37} (34) or σ^{28} (35) have been found in the *spo0B* gene.

Studies on the *spo0B* gene transcription, by detection of mRNA hybridizing to specific probes (results not shown here) and by the analysis of a chromosomal *spo0B-lacZ* fusion have shown that the *spo0B* gene is expressed during vegetative growth and that its expression is slowly decreasing during sporulation. This is in agreement with the observations made by others suggesting that the *spo0B* product is present during vegetative growth (36). The decrease of the *spo0B* gene expression after the end of growth seems to be sporulation-related, as shown by results obtained in a Spo⁻ strain. A different mechanism of regulation seems to control the expression of the *spo0H* gene in *B. licheniformis* (10). These differences could reflect specific temporal roles in the *spo0B* and *spo0H* genes.

It is interesting to note that the expression of the *spo0B-lacZ* fused gene is about 1 order of magnitude lower than that described for the *spoVG-lacZ* gene during vegetative growth, i.e., under "conditions of nonexpression" (22). This low level of expression may reflect the weakness of the *spo0B* promoter due to a -35 region that is distant from the consensus sequence and an insufficient interaction of *spo0B* mRNA with ribosomes, since stronger Shine-Dalgarno pairing is usually found in Gram-positive bacteria (17). Although such a low level of expression may suggest a regulatory role for the *spo0B* gene, it cannot at present be excluded that the *spo0B* gene product (and maybe the products of some other *spo0* genes) takes part in the formation of a spore component. These possibilities are currently investigated in our laboratory.

We would like to thank G. Henckes for helpful advice, Mme. F. de la Torre for skillful technical assistance, and Mme. J. Mauger for excellent typing of the manuscript. This work was supported by a grant from the Centre National de la Recherche Scientifique (LA136).

1. Henner, D. J. & Hoch, J. A. (1980) *Microbiol. Rev.* **44**, 57–82.
2. Piggot, P. J. & Coote, J. G. (1976) *Bact. Rev.* **40**, 908–962.
3. Kawamura, F., Shimotsu, H., Saito, H., Hirochika, H. & Kobayashi, Y. (1981) in *Sporulation and Germination*, eds. Levinson, H. S., Sonenshein, A. L. & Tipper, D. J. (American Society for Microbiology, Washington, DC), pp 109–113.
4. Dubnau, E., Ramakrishna, N., Cabane, K. & Smith, I. (1981) *J. Bacteriol.* **147**, 622–632.
5. Jayaraman, K., Keryer, E. & Szulmajster, J. (1981) *FEMS Lett.* **10**, 273–276.
6. Hirochika, H., Kobayashi, Y., Kawamura, F. & Saito, H. (1981) *J. Bacteriol.* **146**, 494–505.
7. Ferrari, F. A., Lang, D., Ferrari, E. & Hoch, J. (1982) *J. Bacteriol.* **152**, 809–814.
8. Bonamy, C. & Szulmajster, J. (1982) *Mol. Gen. Genet.* **188**, 202–210.
9. Shimotsu, H., Kawamura, F., Kobayashi, Y. & Saito, H. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 658–662.
10. Ramakrishna, N., Dubnau, E. & Smith, I. (1984) *Nucleic Acids Res.* **12**, 1779–1790.

11. Casadaban, M. (1976) *J. Mol. Biol.* **104**, 541-555.
12. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
13. Schaeffer, P., Millet, J. & Aubert, J. (1965) *Proc. Natl. Acad. Sci. USA* **54**, 704-711.
14. Stragier, P., Richaud, F., Borne, F. & Patte, J. C. (1983) *J. Mol. Biol.* **168**, 307-320.
15. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560.
16. Sanger, F. & Coulson, A. R. (1978) *FEBS Lett.* **87**, 107-110.
17. Débarbouillé, M. & Raibaud, O. (1983) *J. Bacteriol.* **153**, 1221-1227.
18. Gilman, M. Z. & Chamberlin, M. J. (1983) *Cell* **35**, 285-293.
19. Shapira, S. K., Chou, J., Richaud, F. & Casadaban, M. (1983) *Gene* **25**, 71-82.
20. Donnelly, C. E. & Sonenshein, A. L. (1982) in *Molecular Cloning and Gene Regulation in Bacilli*, eds. Ganesan, A. T., Chang, S. & Hoch, J. A. (Academic, New York), pp. 63-72.
21. Horinouchi, S. & Weisblum, B. (1982) *J. Bacteriol.* **150**, 815-825.
22. Zuber, P. & Losick, R. (1983) *Cell* **35**, 275-283.
23. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342-1346.
24. McLaughlin, J. R., Murray, C. L. & Rabinowitz, J. C. (1981) *J. Biol. Chem.* **256**, 11283-11291.
25. Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Grothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246**, 40-41.
26. Sollner-Webb, B. & Reeder, R. H. (1979) *Cell* **18**, 485-499.
27. Moran, C. P., Jr., Lang, N., Le Grice, S. F. J., Lee, G., Stephens, M., Sonenshein, A. L., Pero, J. & Losick, R. (1982) *Mol. Gen. Genet.* **186**, 339-346.
28. Rosenberg, M. & Court, D. (1980) *Annu. Rev. Genet.* **13**, 319-353.
29. Hoch, J. A., Shiflett, M. A., Trowsdale, J. & Chen, S. M. H. (1978) in *Spores VIII*, eds. Chambliss, G. & Vary, J. C. (American Society for Microbiology, Washington, DC), pp. 127-130.
30. Losick, R. (1981) in *Sporulation and Germination*, eds. Levinson, H. S., Sonenshein, A. L. & Tipper, D. J. (American Society for Microbiology, Washington, DC), pp. 48-56.
31. Trowsdale, J., Shiflett, M. & Hoch, J. A. (1978) *Nature (London)* **272**, 179-181.
32. Losick, R. & Pero, J. (1981) *Cell* **25**, 582-584.
33. Mulligan, M. E., Hawley, D. K., Enriken, R. & McClure, W. R. (1984) *Nucleic Acids Res.* **12**, 789-800.
34. Wong, S. L., Price, C. W., Goldfarb, D. S. & Doi, R. H. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1184-1188.
35. Gilman, M. Z., Wiggs, J. L. & Chamberlin, M. J. (1981) *Nucleic Acids Res.* **9**, 5991-6000.
36. Brehm, S. P., Le Hegarat, F. & Hoch, J. A. (1975) *J. Bacteriol.* **124**, 977-984.