Nucleotide sequence of the spoB 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, 91900 Gif-sur-Yvette, France

Communicated by Jesse C. Rabinowitz, July 30, 1984

ABSTRACT The spoB 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 spoB gene. The promoter sequence of this gene shows the canonical T-A-T-A-T region at 10 bp from the transcriptional start (−10 region) but an unusual sequence, T-T-T-T-C-M, in the −35 region. The nucleotide sequence shows an open reading frame encoding a 192-amino-acid polypeptide of M, 22,542, which is close to the molecular weight of the spoB product synthesized in E. coli minicells. To investigate the regulation of the spoB gene under a variety of physiological conditions, we constructed an in-frame fusion between the spoB 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 spoB 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 (spoA, spoII, spoIII, spoIV, and spoV) according to the stage of sporulation they are affecting (1). Any mutation in the spo 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 spoA, spoB, etc. (2).

A number of methods recently have been developed for cloning spore genes and applied by different investigators for the isolation of several spo genes (3–7). We have reported the cloning and expression in Escherichia coli of the spoB 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 spoB 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 spoB 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, spoF from B. subtilis (9) and spoH from Bacillus licheniformis (10). It also was of interest to study the expression of the spoB 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 spoB 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 μg/ml, and chloramphenicol, 3 μg/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 Rabaud (17). B. subtilis mRNA was extracted and purified essentially as described by Gilman and Chamberlin (18).

Construction of a spoB-lacZ Fusion. pSKS107 plasmid (19) was first deleted after partial AvaI digestion, which eliminated the distal part of the lac operon, known to be toxic in B. subtilis cells (20). A 915-base-pair (bp) HindIII fragment was purified from pGsoB10 plasmid (8). This fragment carried the HindIII–BamHI part of pBR322 and the first 590 bp of the sequence given in Fig. 3. Insertion of this fragment in the unique HindIII site of pSKS107A AvaI led to an in-frame fusion between spoB and lacZ coding sequences (plasmid pOB108). A 1135-bp Nae I–Tag I fragment carrying a chloramphenicol-resistance (Cm') determinant was purified from plasmid PC194 (ordinate 975–2110 in ref. 21) and inserted between the SmaI and SalI 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 spoB gene stretches

Abbreviations: bp, base pair(s); kb, kilobase(s); Cm', chloramphenicol resistance.
Fig. 1. Sequencing strategy of the spoOB 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 spoOB gene.

On both sides of the unique Pst I site at the position 567 (7) and runs over a Sau3A site located between HindIII 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 spoOB 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 spoOB 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 Mr 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 spo genes, spoOF from B. subtilis (9) and spoOH from B. licheniformis (10). There is also no homology to any other known proteins.

Transcriptional Start Site of the spoOB Gene. The promoter region of the spoOB 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 HindIII–Hinfl 67-bp fragment or the HindIII–Taq I 432-bp fragment labeled at position 432

![Diagram](image)

**Biochemistry: Bouvier et al.**


---

**Fig. 2.** Nucleotide sequence of the spoOB 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 spoOB mRNA start; the corresponding −10" and −35" promoter signals are underlined with a thick line. The HindIII site used in the course of the construction of a hybrid spoOB-lacZ gene is shown.
their HindIII 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 fragment generated by the chemical sequence analysis performed on the HindIII–Taq 1 probe. Comparison of the results obtained with both methods allowed unambiguous localization of the spoOB 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 σ70 subunit (27). This region is separated by 17 bp from a T-T-T-C-T-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 spoOB 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 (∆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 spoOB-lacZ Fusion. To study the regulation of the expression of the spoOB 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 spoOB 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 spoOB that allows the synthesis of a hybrid protein of which the first 45 amino acids are encoded by the spoOB gene and which is endowed with β-galactosidase activity. This plasmid carries also the Cm' 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' 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 spoOB gene as detailed in ref. 22. This was first verified by cotransformation of the Cm' marker with the phe marker

![Fig. 3](https://example.com/fig3.png)

**Fig. 3.** Mapping of the 5' end of the spoOB transcript. After hybridization with vegetative mRNA from strain B31/psOB10 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](https://example.com/fig4.png)

**Fig. 4.** Physical map of plasmid pOB110 carrying a spoOB-lacZ fusion. Arrows show the direction of transcription of the Cm' and ampicillin (Ap)-resistance genes and of the hybrid spoOB-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 spoOB DNA. The Sma I and Nae I sites are not regenerated.

![Fig. 5](https://example.com/fig5.png)

**Fig. 5.** β-Galactosidase activity in the chromosomal spoOB-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: B520 grown in DS medium (●) and in DS medium with 0.5% glucose (○) and B31 (spoOB) grown in DS medium (●). Results given are the average of three experiments.
closely linked to the spoOB gene on the B. subtilis chromosome (data not shown). Moreover, when using the asporogenic strain B331 carrying the 0B136 mutation localized in the proximal part of the spoOB gene (7), both Spom and Spoc Mm colonies were recovered, confirming integration of the pOB110 plasmid into the spoOB region. In that case, Spom 
Cm' 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 spoOB gene is maximal at the beginning of the stationary phase, which progressively restored the level 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 transient 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 spoOB 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 spoOB 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 that obtained in E. coli minicells. Our result is consistent with the 39-kDa polypeptide reported by Hirochika et al. (6) and assumed to be encoded by an EcorI fragment carrying the spoOB gene cloned into the transducing phage ρ11. The size of the spoOB gene product estimated from our current study is similar to those encoded by the spoOF gene of B. subtilis (173 amino acids (9)) or by the spoOH gene of B. licheniformis (168 amino acids (10)).

The knowledge of the primary structure of the sequenced spoOB gene 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 spoO genes are required for the progression of sporulation events (29, 30). Concerning specifically the spoOA and spoOB 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 spoOA and spoOB 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 spoOB 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 factors associated with the B. subtilis RNA polymerase (32). The promoter recognized by the major form of RNA polymerase containing σ21 is similar to that found in E. coli. The promoter found in the spoOB gene is apparently of the same nature; it contains the canonical sequence T-A-T-A-T-A 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-T-T sequence located at 17 bp from the −10 region or a T-T-G-T-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 factors designated σ70 (34) or σ21 (35) have been found in the spoOB gene.

Studies on the spoOB gene transcription, by detection of mRNA hybridizing to specific probes (results not shown here) and by the analysis of a chromosomal spoOB-lacZ fusion have shown that the spoOB 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 spoOB product is present during vegetative growth (36). The decrease of the spoOB gene expression after the end of growth seems to be sporulation-related, as shown by results obtained in a Spom strain. A different mechanism of regulation seems to control the expression of the spoOH gene in B. licheniformis (10). These differences could reflect specific temporal roles in the spoOB and spoOH genes.

It is interesting to note that the expression of the spoOB-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 spoOB promoter due to a −35 region that is distant from the consensus sequence and an insufficient interaction of spoOB 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 spoOB gene, it cannot at present be excluded that the spoOB gene product (and maybe the products of some other spoO 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).