Identification of the promoter and the transcriptional start site of the spoVA operon of Bacillus subtilis and Bacillus licheniformis

B. MOLDOVER,¹ P. J. PIGGOT¹ and M. D. YUDKIN²*

¹Department of Microbiology, Temple University School of Medicine, Philadelphia, PA 19140, USA
²Microbiology Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

(Received 17 July 1990; revised 27 August 1990; accepted 19 September 1990)

The region upstream of the coding sequence of the spoVA operon was studied by several techniques to identify the promoter and to determine the start point for transcription of spoVA. The results of plasmid integration analysis in Bacillus subtilis showed that no more than 119 bp upstream of the coding sequence is needed for expression. A comparison of the sequence of this upstream region with the corresponding sequence from Bacillus licheniformis showed four stretches that were perfectly conserved, interspersed with poorly conserved stretches; the second and third of these conserved stretches appeared to represent the '−35' and '−10' regions of a promoter recognized by RNA polymerase containing σG. Primer extension analysis in B. subtilis revealed a spoVA transcript which had apparently initiated 6 bp downstream of the putative '−10' heptanucleotide CATACTA, that is, 27 bp upstream of the coding sequence. This transcript was observed 4 h and 5 h after the initiation of sporulation, but not at earlier times.

Introduction

The formation of endospores in bacilli requires the temporally regulated expression of several dozen genes. One of the mechanisms that ensures the correct regulation of expression is the sequential synthesis of a number of sigma factors, which interact successively with the core RNA polymerase. Each sigma factor, in combination with core RNA polymerase, recognizes its cognate promoters and hence leads to transcription of a specific set of genes, and each set of promoters is presumed to differ in nucleotide sequence from the other sets.

One genetic locus that regulates sporulation is spoIIA, which both of our laboratories have studied extensively. The spoIIA locus comprises three genes (Fort & Piggot, 1984; Piggot et al., 1984), and is expressed in Bacillus subtilis within an hour of the onset of sporulation, from a promoter that resembles promoters recognized by RNA polymerase containing σH (Wu et al., 1989). In B. licheniformis, the three genes of spoIIA are identical in length, and very similar in sequence, to those of B. subtilis, and the putative promoter is precisely conserved between the two species (Yudkin et al., 1989). About 120bp downstream of spoIIA in B. subtilis lies another sporulation locus, spoVA. There are several reasons for believing that in B. subtilis spoVA is transcribed separately from spoIIA: first, experiments with integrational plasmid vectors suggest that the two loci belong to different transcriptional units (Piggot et al., 1984); secondly, the sequence immediately downstream of spoIIA includes what appears to be a terminator of transcription (Wu et al., 1989; Yudkin et al., 1989); thirdly, results of experiments with fusions of spoVA to lacZ show that spoVA is expressed much later in sporulation than spoIIA (Errington & Mandelstam, 1986); fourthly, expression of spoVA, unlike that of spoIIA, must take place in, and is confined to, the prespore compartment (Errington & Mandelstam, 1986; Gholamhoseinian & Piggot, 1989); fifthly, the size of spoIIA mRNA estimated from Northern blots agrees with that predicted for a tricistronic spoIIA message (Savva & Mandelstam, 1986). We now describe experiments whose results have identified the startpoint of transcription of spoVA, and show that the spoVA promoter, which is identical in sequence in B. subtilis and B. licheniformis, resembles promoters that are known to be recognized by RNA polymerase containing σG.
Methods

Bacterial strains. The Escherichia coli strain used was DH5α [F− endA1 hisD17 (rpsL::Tn10) supE44 thi-1 lacI− recA1 gyrA96 relA1 (lacZYA-argF)U169 80lacZAM15]. The B. subtilis 168 strain used was MB24 (trpC2 metC3 rfp-). B. subtilis strains containing a single copy of an integrative plasmid were constructed by transformation with a plasmid containing the chloramphenicol drug resistance marker, and maintained on media containing 3-5 µg chloramphenicol per ml⁻¹.

Plasmids. Plasmids were maintained in E. coli DH5α unless otherwise noted. Plasmid pPP151 contains a 952 bp fragment covering argF169 980dlacZAM151. The B. subtilis strains containing a single copy of an integrative plasmid were constructed by transformation with a plasmid containing the chloramphenicol drug resistance marker, and maintained on media containing 3-5 µg chloramphenicol per ml⁻¹.

DNA preparation. Plasmid DNA was prepared from E. coli by the method of Birnboim & Doly (1979), and, for small-scale preparation, by the method of Ish-Horowicz & Burke (1981).

RNA preparation. RNA from B. subtilis was prepared as previously described (Wu et al., 1989).

Primer extension. Primer extension analysis was by the method of Inoue & Cech (1985). Synthetic primers were labelled at the 5' ends by incubation with T4 polynucleotide kinase in the presence of α-32P]ATP at 37 °C for 60 min. 10 pg labelled primer was mixed with 100 µg RNA in a final volume of 20 µl hybridization buffer (0.2 M-KCl, 0.1 M-Tris/HCl, pH 8.3). The mixture was first heated to 90 °C, and then incubated at 55 °C for 10 min, and on ice for 5 min, to allow for DNA-RNA hybridization. The mixture was then precipitated with 2.5 vol. ethanol, dried, and resuspended in 9 µl reverse transcription buffer (50 mM-Tris/HCl, pH 8.3, 50 mM-KCl, 6 mM-MgCl₂, 1 mM-dithiothreitol, 1 mM-dNTP's, 2 units RNAsin). To this was added 3-5 µl of reverse transcription buffer (50 mM-Tris/HCl, pH 8.3, 50 mM-KCl, 6 mM-MgCl₂, 1 mM-dithiothreitol, 1 mM-dNTP's, 2 units RNAsin). This was then heated at 75 °C for 2 min, and 3 µl was loaded on an 8% (w/v) urea-8% (w/v) polyacrylamide sequencing gel for analysis. The same primer was used for DNA sequencing reactions with pPP171 as template, and the products were loaded on the same gels.

Other methods are described by Wu et al. (1989).

Results

Previous experiments with integrative plasmids had suggested that spoVA is in a different transcriptional unit from spoIIA (Piggot et al., 1984). We therefore delimited the 5' end of the spoVA transcriptional unit of B. subtilis by the use of a new integrational plasmid. Plasmid pPP170 contains the 2250 bp Ps1I fragment that extends from 119 bp upstream of the spoVA coding sequence to the middle of the fourth ORF in spoVA. Chloramphenicol-resistant transformants of B. subtilis MB24 were obtained in which pPP170 had integrated by a Campbell-like mechanism (as confirmed by Southern analysis of appropriately restricted chromosomal DNA). These transformants sporulated normally. This indicated that the cloned region in pPP170 extends outside the spoVA transcription unit, indicating that no more than 119 bp upstream of the coding sequence is needed for expression of spoVA. We then examined this upstream region.

Yudkin et al. (1989) isolated a clone from B. licheniformis that contained the spoIIA region. Sequencing showed that the third gene of the region was succeeded by a sequence that resembled a strong transcription terminator. We have now obtained the sequence of a further 150 bp downstream of spoIIA: this sequence is compared with the corresponding region of B. subtilis in Fig. 1. We note the following features.

The last 51 nucleotides shown begin with ATG and represent an open reading frame whose putative translation product is extremely similar to that at the start of spoVA in B. subtilis. Given this similarity, and the fact that, as in B. subtilis, the open reading frame begins about 100 bp downstream of spoIIA, we assume that we have sequenced the beginning of the spoVA region of B. licheniformis. Five to 14 bp upstream of the B. licheniformis open reading frame lies a sequence whose transcript could function as a ribosome binding site: base-pairing with the 3' end of 16S RNA would give a structure with free energy of about −14.5 kcal mol⁻¹ (60.1 kJ mol⁻¹) according to the rules of Tinoco et al. (1973).

Between the apparent terminator of spoIIA and this putative ribosome binding site lie four stretches of sequence that are perfectly conserved within the two species, interspersed with stretches in which the two sequences differ. Since we presume that the spoVA promoter will lie in this region, we wondered whether a '−35' and a '−10' sequence could be found among the four conserved stretches.

We therefore determined the initiation site of transcription of spoVA in B. subtilis by primer extension analysis. We extracted RNA from cultures during exponential growth, at the end of exponential growth (which we define as T₀, the point at which sporulation is induced), and hourly thereafter. Samples of this RNA were annealed to two primers identical in sequence to 17-base sections of the transcribed strand of the DNA, one representing bases 90 to 74, and the other bases 32 to 16 of the published sequence (Fort & Errington, 1985; their numbering is used); these were then used as primer templates for the AMV reverse transcriptase. The same primers were used to initiate dideoxy sequencing reactions (Sanger et al., 1977) with pPP171 as template, and the products of the different enzyme reactions from each primer were fractionated on urea-polyacrylamide gels.

The results (Fig. 2) show that a transcript corresponding to spoVA was present at T₄ and T₅, but not in vegetative growth; we also found no transcript at T₀, T₁, T₂, or T₃ (results not shown). [In spoVA : : lacZ fusions...
Transcriptional start site of the spoVA operon

I

GTGGTAATTTATGGTCTTTTGCCTTTCTGAATGATAAATGGGATGTACTTT
GTGGTAATTTATGGTCTTTTGCCTTTCTGAATGATAAATGGGATGTACTTT

---

II

MER
MER

III

IV

CATACTACAACCTATAACCATCATATAGGAAGTGACCCA.GATGGAACGTCAA
CATACTACAACCTATAACCACCACCACCAGAGTATGGATGGAACGACGA

---

IFIRHLRHRVLAHP
VFIRLRRHRLHEADP
GTTTTATCAAGACTCCGCACCCGGCTTGGAGGCAGATCCG
ATATTTATCCGCTTGGACCCGAGTGCGTGCACATCCA

Fig. 1. The sequence downstream of the terminator of spoIIA. The nucleotide sequence in heavy type is that of Bacillus licheniformis; the first nucleotide shown is immediately 3' of the sequence given in Fig. 5 of Yudkin et al. (1989). The amino-acid sequence shown (in heavy type) immediately above the nucleotide sequence is that of the predicted N-terminal region of the product of the spoVA gene. The nucleotide and amino-acid sequences in normal type are those reported for the corresponding region of Bacillus subtilis (Fort & Piggot, 1984; Fort & Errington, 1985). Dots mark the position of gaps that have been introduced to maximize the homology. The dashed underlining and Roman numerals indicate the regions of nucleotide conservation between the two species of Bacillus.

\[ \beta\text{-galactosidase activity is detectable at } T_3 \text{ and much increased at } T_4 \text{ and } T_5 \text{ (Errington & Mandelstam, 1986).} \]

Experiments with each primer clearly showed only a single transcript, which appeared from its length to have its 5' end at the second nucleotide in region IV shown in Fig. 1, which corresponds to nucleotide 2051 of the published sequence.

In the course of the sequencing reactions, we detected a G (open arrow in Fig. 2) which is not present in the published sequence (Fort & Piggot, 1984).

Discussion

The results show unequivocally that in B. subtilis spoVA – or, to be more precise, the first gene of spoVA – lies in a different transcriptional unit from spoIIA. Not only are the two loci expressed at different times during sporulation, but the starting point of the transcript that includes spoVAA is well downstream of the terminator that

---

Fig. 2. Determination of the 5' end of the spoVA transcript of B. subtilis. Primer extension analysis was carried out on mRNA prepared as described in Methods. RNA was extracted from M24 during vegetative growth (sample V) and during sporulation at T4 (sample 4) and T5 (sample 5). The primer used was identical in sequence to bases 90 to 74 of the published sequence (Fort & Errington, 1985). A sequencing ladder using the same primer is also shown. The letters above the lanes indicate which deoxy nucleotide was used to terminate the sequencing reaction. The putative transcription start site is indicated with a solid arrow. The open arrow indicates the extra G found during sequencing.
follows the final open reading frame (spolIAC) of the spolIIA locus. The sequence of spolVA gives no indication that the five genes of the locus comprise more than one transcription unit, and experiments with integrative plasmids also suggest that the five are expressed from a single promoter (Fort & Errington, 1985), although the results of Northern blots were not definitive on this point (Savva & Mandelstam, 1986).

Formally, we have identified the 5' end of the spolVA mRNA. We consider this to be the transcriptional start point used in vivo, as it agrees exactly with the transcription start point obtained in vitro with purified E-σ^70 (Sun et al., 1989), and there is no evidence of mRNA processing. The identification of the starting point for transcription of spolVA allows us to conclude that the sequences TGAATG and CATACTA (regions II and III in Fig. 1) are the '−35' and '−10' recognition sequences of the spolVA promoter and that they are separated by a 17 bp 'spacer'. The 5' nucleotide of the mRNA then corresponds to the A located six nucleotides downstream of the CATACTA sequence. These recognition sequences correspond exactly to the consensus sequences of σ^70 promoters; and the conclusion that spolVA is transcribed by RNA polymerase containing σ^70 is consistent with the fact that this form of the enzyme interacts specifically with one of the several available sigma factors.

Two other regions (I and IV, Fig. 1) are strongly conserved in addition to regions II and III which are the '−35' and '−10' regions. Region IV includes the transcription start point of spolVA and the first 7 nucleotides of the transcribed region. The striking conservation of the 20 bp region I, extending from 45 to 64 bp upstream of the deduced spolVA transcription start point, leads us to think that the region may be important for the regulation of spolVA expression. No homology was detected between region I and the corresponding upstream regions of genes gerA (Feavers et al., 1985), sspA, sspB (Conners et al., 1986), sspC (Conners & Setlow, 1985), sspD (Conners et al., 1986), and sspE (Hackett & Setlow, 1987). In contrast, the '−35' and '−10' regions of these genes share extensive homology with the '−35' and '−10' regions of spolVA, and indeed are transcribed by E-σ^70. Thus region I may be a control region that is specific to spolVA.

B. M. and P. J. P. were supported in part by Public Health Service Grant GM43577 from the National Institute of Health and by grant DMB-8912323 from the National Science Foundation. We are grateful to Miss L. Appleby for technical help.

References


