A novel member of the subtilisin-like protease family from *Bacillus subtilis*

Angela Valbuzzi,¹ Eugenio Ferrari² and Alessandra M. Albertini¹

Author for correspondence: Alessandra M. Albertini. Tel: +39 0382 505549. Fax: +39 0382 528496.
e-mail: albert@pillo.unipv.it

¹ Dipartimento di Genetica e Microbiologia, Università degli Studi di Pavia, Via Abbiategrasso, 207, 27100 Pavia, Italy

² Genencor International, 925 Page Mill Road, Palo Alto, CA 94304, USA

***aprX* is a 1326 bp gene of *Bacillus subtilis* strain 168 that encodes a serine protease, probably intracellular, characterized by significant similarity with subtilisins, thermitases and pyrolysins. Transcription analysis, performed by RT-PCR and primer extension, allowed the localization of the active promoter and showed that *aprX* is expressed in stationary phase. The pattern of expression of *aprX* and its dependence on various transition state regulatory genes (*degU*, *degQ*, *hpr*, *abrB*, *sinR*), monitored by *lacZ* transcriptional fusions, are distinctive from those of subtilisin and other degradative enzymes. *aprX* is not essential for either growth or sporulation.**

Keywords: *Bacillus subtilis*, subtilase, stationary phase, transcription analysis

INTRODUCTION

Many species of the genus *Bacillus* produce a variety of extracellular and intracellular proteases (Pero & Sloma, 1993). The major extracellular proteolytic enzymes are neutral (metallo-) and alkaline serine proteases, such as subtilisin produced by *Bacillus subtilis*. Subtilisin is one of the most extensively studied of all bacterial proteins, due to its commercial importance (Harwood, 1992; Wells & Estell, 1988). This protease is produced at the onset of sporulation, and its expression is controlled by the complex network regulating the transition state, including the action of several *spo0* genes, *abrB*, *sinR*, *hpr*, *degU* and other regulators (Ferrari *et al.*, 1993). Subtilisin is synthesized as a pre-pro-protein, translocated through the cell membrane via the pre-peptide (or signal peptide) and finally activated by auto-proteolytic removal of the pro-peptide (Power *et al.*, 1986), which functions as an intramolecular chaperone (Shinde *et al.*, 1997). This protease is not essential for either growth or sporulation, but is probably used in nature as a scavenging enzyme.

At present more than 200 proteases have been assigned to the superfamily of subtilases (subtilisin-like serine proteases), with representatives both in micro-organisms (archaea, bacteria, fungi and yeast) and in higher eukaryotes (Siezen & Leunissen, 1997). All the enzymes belonging to this superfamily have in common a core structure, the catalytic domain, characterized by the presence of 'structurally conserved regions', which correspond to common secondary structure elements. Most of the subtilases characterized so far are extracellular and are subdivided into six families, according

to the sequence similarity. The subtilisin family includes true subtilisins as well as minor intracellular proteases (Siezen & Leunissen, 1997). Some of the intracellular serine proteases from *B. subtilis*, such as ISP-I (Koide *et al.*, 1986), also called IspA, play an important role in protein turnover or processing during sporulation or may be involved in the heat-shock response (HtrA or serine protease Do; Devine & Noone, 1998).

We report here the identification and partial characterization of the *aprX* gene, whose product reveals high similarity with subtilisins from different *Bacillus* species. To investigate the physiological role of *aprX*, we studied the regulation of its transcription.

METHODS

Bacterial strains, plasmids and media. The *Bacillus subtilis* strains used in this study are listed in Table 1. *Escherichia coli* DH5 α (*supE44 lacU169 ϕ 80 lacZ Δ M-15 *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*) was used as a host for plasmid constructions.*

The integrative vectors pJM783 (Perego, 1993) and pJM115, a derivative of pDH32 (Perego, 1993), were used to obtain transcriptional fusions to the *lacZ* gene. PCR products of different length were amplified from the *aprX* upstream region using primers with *EcoRI*–*Bam*HI sites. The primer L378 (nt 378–358 in Fig. 1 and 80640–660 in EMBL sequence no. Z99113), internal to the coding sequence, was used in combination with one of the following upstream primers, U147 (nt 147–167 in Fig. 1 and 80871–851 in EMBL sequence no. Z99113), 5'-GGAATTCGTTTTGCTTTTCTGTCTCA-3' (nt 81095–074 in EMBL sequence no. Z99113) and 5'-GGAATTCATGAGGTTTAGCCAATAGGT-3' (nt 81453–433 in EMBL sequence no. Z99113), to generate,

Table 1. *B. subtilis* strains used in this study

Strain	Genotype	Reference/source
168	<i>trpC2</i>	Laboratory stock
BFS2601	<i>trpC2 aprX::pMUTIN4/298HB erm</i>	This study
BG2812	$\Delta npr hpr-2$	E. Ferrari
BG2817	$\Delta npr degU^{hy}-32$	E. Ferrari
BG2821	$\Delta npr degQ^{hy}$	E. Ferrari
BG2822	Δnpr	E. Ferrari
PB1768	<i>hisA1 thr-5 trpC2</i>	R. Dedonder, Institut Pasteur, Paris, France
PB1881	<i>hisA1 thr-5 trpC2 amyE::pJM115/232EB aprX-lacZ kan</i>	This study
PB1882	<i>hisA1 thr-5 trpC2 amyE::pJM115/455EB aprX-lacZ kan</i>	This study
PB1883	<i>hisA1 thr-5 trpC2 aprX::pJM783/232EB lacZ cat</i>	This study
PB1884	<i>hisA1 thr-5 trpC2 aprX::pJM783/455EB lacZ cat</i>	This study
PB1885	$\Delta npr hpr-2 aprX::pJM783/455EB lacZ cat$	This study
PB1886	$\Delta npr degU^{hy}-32 aprX::pJM783/455EB lacZ cat$	This study
PB1887	$\Delta npr degQ^{hy} aprX::pJM783/455EB lacZ cat$	This study
PB1888	$\Delta npr aprX::pJM783/455EB lacZ cat$	This study
PB1889	<i>hisA1 thr-5 trpC2 $\Delta sin::phleo aprX::pJM783/455EB lacZ cat$</i>	This study
PB1891	<i>hisA1 thr-5 trpC2 abrB::Tn917erm aprX::pJM783/455EB lacZ cat</i>	This study
PB1894	<i>hisA1 thr-5 trpC2 amyE::pJM115/814EB aprX-lacZ kan</i>	This study
PB1906	<i>trpC2 $\Delta aprX::cat$</i>	This study
PB5090	<i>leuA8 metB5 hisA1 $\Delta sin::phleo$</i>	I. Smith, The Public Health Research Institute, New York, USA
PB5101	<i>trpC2 pheA1 abrB::Tn917erm pksX::pJM783/1197EB cat</i>	C. Scotti, DGM, Università di Pavia, Italy

respectively, 232, 455 and 814 bp PCR products. The purified products were cloned into pJM783 and/or pJM115.

A 298 bp fragment from nt 80529 to nt 80232 in EMBL sequence accession number Z99113, obtained by PCR, was cloned in plasmid pMUTIN4 (Vagner *et al.*, 1998) to perform *aprX* inactivation by single crossover insertion. Fragments of 315 bp and 243 bp from nt 80972 to nt 80658 and from nt 79651 to nt 79409, respectively, obtained by PCR, were cloned in plasmid pJM105A (Perego, 1993) to perform *aprX* disruption by double crossover insertion. *E. coli* DH5 α was grown in LB medium. When required, ampicillin was added at a concentration of 100 μ g ml⁻¹. *B. subtilis* strains were grown in NB medium (Difco), supplemented, when required, with 2.5 or 5 μ g chloramphenicol ml⁻¹, 2 μ g kanamycin ml⁻¹ or 1 μ g erythromycin ml⁻¹.

DNA manipulations. Standard DNA manipulations were carried out as described by Sambrook *et al.* (1989). *B. subtilis* chromosomal DNA was prepared by phenol extraction (Albertini & Galizzi, 1985). DNA sequencing on both strands was conducted by the dideoxy chain-termination method with the T7 (Pharmacia) or Thermo Sequenase (USB) sequencing kits.

RNA extraction, RT-PCR and primer extension analysis. Samples (0.3–3 ml) of PB1768 parental strain cells growing in 2 $\times$ NB (Difco) medium were collected at various times. Total RNA was extracted using the Qiagen 'RNeasy mini' or 'midi' kit, incubating the cell suspension at 37 °C for 20 min during the lysis step. All the RNA samples were treated with DNase I (RNase free) (Boehringer) and then repurified with the 'RNeasy clean-up protocol' (Qiagen). The absence of contaminating DNA was tested by performing PCR reactions on

RNA with 15 pmol reverse primer L468 (annealing at nt 468–450 in Fig. 1, corresponding to nt 80550–568 in EMBL sequence no. Z99113) and of forward primer 5'-TATGAG-GGTTTAGCCAATAGGT-3' (negative control primer, annealing at nt 81453–433 of EMBL sequence no. Z99113). Reverse transcription (RT) experiments were performed with 2 μ g total RNA, according to the Promega Primer Extension System, using the quoted reverse primer. Following the primer extension reaction, we incubated the samples for 2 min at 94 °C. PCR reactions were performed adding to one-third of the samples 1 U *Taq* polymerase (Pharmacia), PCR buffer and 15 pmol of one of the following alternative forward primers: negative control primer (as above), positive control primer U262, annealing at nt 262–281 in Fig. 1 (nt 80756–737 of EMBL sequence no. Z99113), and discriminating primer U95, annealing at nt 95–115 in Fig. 1 (nt 80923–903 of EMBL sequence no. Z99113).

The primer extension experiment was performed according to the Promega Primer Extension System with total RNA extracted from PB1768 cells grown in 2 $\times$ NB medium and harvested at t_0 . Sixteen picomoles of the primer L378 (Fig. 1) was radiolabelled by T4 polynucleotide kinase using 30 μ Ci (1.11 MBq) [γ -³²P]ATP (> 5000 Ci mmol⁻¹; Amersham). Primer extension was performed using 1.6 pmol labelled L378 and 100 μ g total RNA.

β -Galactosidase assays. β -Galactosidase activity in *B. subtilis* strains harbouring *aprX-lacZ* transcriptional fusions was measured as described previously (Scotti *et al.*, 1996). The cultures, aerated by shaking, were grown at 37 °C in TM (Anagnostopoulos & Spizizen, 1961) minimal medium or 2 $\times$ NB (Difco) medium.

Sporulation assays. The sporulation frequencies of *B. subtilis* strains were measured in samples of liquid cultures grown at 37 °C in Schaeffer medium (Schaeffer *et al.*, 1965) at t_{16} and t_{22} in the late-stationary phase. Viable spores were counted by plating on NB medium suitable dilutions of 0.5 ml samples treated for 10 min at room temperature with 50 µl chloroform; the titre was compared to that of untreated samples.

RESULTS AND DISCUSSION

Nucleotide sequence and deduced product analysis of the *aprX* gene

During the genome sequencing project (Kunst *et al.*, 1997) we isolated, via plasmid rescue in the genome region around 158.9°, pDA11 and pAA12, derivatives of the integrative vectors pDIA5304 (Glaser *et al.*, 1993) and pJM103 (Perego, 1993), respectively, bearing two overlapping fragments covering the region from *pksS* to *miaA*. Nucleotide sequence analysis of the region, part of the EMBL accession number Z99113, revealed a 1326 bp ORF (*aprX*) followed by a short ORF of 699 bp (*ymaE*), both in reverse orientation with respect to the direction of replication (Kunst *et al.*, 1997). *aprX* is preceded by two putative σ^A promoter consensus sequences (P_L and P_S), each consisting of conserved -35 and -10 sequences. Two start codons (ATG and TTG) preceded by a sequence with good consensus to the canonical ribosome-binding site (Fig. 1) are in the same frame, therefore translation could generate two hypothetical gene products, AprX_L and AprX_S, 442 or 417 amino acids long, respectively. Downstream of *aprX* there is a potential hairpin terminator (ΔG -105.1 kJ mol⁻¹), followed by a putative σ^A promoter sequence with a weak -35 consensus, upstream of *ymaE*. Amino acid sequence similarity analyses with the BLASTP (Altschul *et al.*, 1990) and FASTA (Pearson, 1990) programs strongly indicated that the deduced product

AprX is a serine protease of the subtilase superfamily. Dendrogram analysis based on alignment of multiple sequences of the catalytic domain, from residue 112 of AprX_L, emphasizes the relationship of AprX with the families of subtilases sharing highest similarity scores in the FASTA search (Fig. 2). The similarity to subtilases starts from residue 117 of AprX_L and extends to the C-terminus of the protein. The active site residues Asp-155, His-187 and Ser-384 are clearly conserved, as well as the entire catalytic domain, including the 'structurally conserved regions', as defined by Siezen & Leunissen (1997). Within the catalytic domain the peculiar characteristics of a subtilase are obvious even if, unlike the majority of subtilases, this enzyme lacks a canonical signal sequence for membrane translocation (signal peptide). This finding suggests an intracellular location for AprX. However, the low sequence similarity with the intracellular serine proteases of *Bacillus polymyxa* (38.5% identity with BpIsp) and *B. subtilis* (33.3% identity with BsIsp1) does not allow the classification of this subtilase in the sub-group of known intracellular subtilisins. AprX_L contains an N-terminal extension (residues 1-116), relative to the catalytic domain, with no similarity to known proteins. It should also be noted that this new serine protease contains nine Cys residues, four of which are in the catalytic portion. Some of these residues, extremely rare in subtilases from Gram-positive bacteria, are nevertheless conserved in subtilases from other organisms (Siezen & Leunissen, 1997). In conclusion, AprX, although strongly related to subtilases, appears to be a member of a new family.

aprX transcription and regulation

To analyse the transcription of *aprX* we performed RT and PCR experiments (RT-PCR) on total RNA extracted from the wild-type strain PB1768 at various

Fig. 1. Sequence of the *aprX* gene regulatory region. The first nucleotide corresponds to nt 81017 in EMBL sequence accession number Z99113; the last nucleotide corresponds to nt 80503. PS and PL indicate, respectively, the two putative 'short' and 'long' promoter regions; thin lines indicate the consensus sequences to the canonical σ^A promoter. +1 indicates the transcription start site preceded by the promoter sequence (weak -35 and -10 σ^A consensus sequences in bold characters). RBS indicates the positions of the Shine-Dalgarno consensus sequences just upstream of the two putative ATG and TTG start codons (in bold). Primers used for RT-PCR and primer extension experiments are also indicated.

Fig. 2. Family tree of the subtilase families compared to AprX: the CLUSTAL W (Thompson et al., 1994) output (from FASTA search; Pearson, 1990) for phylogenetic tree construction was used by the program NJPLOT (Perrière & Gouy, 1996). We used the following abbreviations to refer to various bacterial and low eukaryotic subtilases (species and SWISS-PROT accession numbers are indicated): BsAprX, *B. subtilis* AprX; Bslsp1, *B. subtilis* P11018; Bplsp, *B. polymyxa* P29139; BsAprE, *B. subtilis* P04189; BmApr, *Bacillus mesentericus* P07518; BnAprN, *B. subtilis* var. *natto* P35835; BasBpn, *Bacillus amyloliquefaciens* P00782; BlsCar, *Bacillus licheniformis* P00780; BsAprDy, *B. subtilis* P00781; BacEly2, *Bacillus* sp. P41363; BsEpr, *B. subtilis* P16396; TvTher, *Thermoactinomyces vulgaris* P04072; AnPepC, *Aspergillus niger* P33295; ScPrb1, *Saccharomyces cerevisiae* P09232; SpSepr, *Schizosaccharomyces pombe* P40903; ScYct5, *S. cerevisiae* P25381; BsBpf, *B. subtilis* P16397.

times. Using the positive control primer U262 we detected a product of 206 bp in samples 2 and 5 (corresponding to t_6 and t_2), while no product was present in t_{-1} and t_0 samples (11 and 8 in Fig. 3). These results indicated that *aprX* is transcribed in rich medium ($2 \times$ NB) from t_2 to t_6 (t_0 represents the time of transition from the exponential to the stationary growth phase; the growth time is measured in hours before or after t_0). The use of alternative forward primers gave us some indications about the transcription start site: no product was present in samples obtained using the discriminating primer U95 (3, 6, 9 and 12). These data suggested that, at the times tested, the transcription start site of *aprX* must be located in the proximity of the σ^A P_S promoter of Fig. 1. Moreover these results ruled out the use of P_L

Fig. 3. Agarose gel electrophoresis of RT-PCR products obtained from total RNA extracted from PB1768 grown in $2 \times$ NB medium. Lanes: M, size markers; 1, 4, 7 and 10, PCR products obtained from RT of RNA extracted at t_6 , t_2 , t_0 and t_{-1} with the forward negative control primer; 2, 5, 8 and 11, PCR products obtained from RT of RNA extracted at t_6 , t_2 , t_0 and t_{-1} with the forward positive control primer U262; 3, 6, 9 and 12, the results of the PCR experiments performed with the forward discriminating primer U95 and RT of RNA extracted at t_6 , t_2 , t_0 and t_{-1} . For all the samples RT was performed as described in the text with the same reverse primer, L468. The arrow indicates the position of the 206 bp product.

as a secondary promoter. The primer extension analysis allowed us to map the transcription start site at nt 238 in Fig. 1 (nt 80780 in EMBL sequence no. Z99113). The start site is in fact located in the P_S region, but does not match the σ^A consensus sequence previously identified in this region. Only a sequence with a moderate σ^A consensus results in being properly located.

To define the minimal region containing the transcription control sequences of *aprX*, fragments of different length, amplified from the putative promoter region of *aprX*, were cloned in the transcriptional fusion vector pJM115. We tested their ability to promote transcription of the *lacZ* gene, after integration of the plasmid by double crossover within the α -amylase gene of the PB1768 strain. The 232 bp fragment contained the proximal P_S region (Fig. 1), the 455 bp fragment also included the distal P_L sequence, whereas the 814 bp fragment corresponded to the entire intercistronic stretch between *ymaC* and *aprX*. None of the corresponding transformed strains (PB1881, PB1882 and PB1894) produced significant levels of β -galactosidase activity [less than 2.0 Miller units (mg protein) $^{-1}$ for all the times tested from t_{-1} to t_6 , data not shown].

To evaluate if the absence of expression was due to a lack of promoter activity or to the particular ectopic position of the fusion, the same 232 and 455 bp fragments were cloned in pJM783. This integrative plasmid allows the insertion of the *lacZ* transcriptional fusion by single crossing-over only into the *aprX* locus. The eventuality of amplification of the constructs was excluded by Southern blot analysis of chromosomal DNA extracted from the transformed colonies, ap-

Table 2. Expression of *aprX* in different genetic backgrounds, measured 2, 4 and 9 h after t_0 (the time of transition from exponential to stationary growth phase)

Results are expressed as β -galactosidase activity [Miller units (mg protein)⁻¹]. Underlined values are the mean values from two different experiments.

Strain*	TM medium			2 × NB medium		
	t_2	t_4	t_9	t_2	t_4	t_9
PB1883 (wt)	0.05	0.1	0.03	3.0	3.2	6.6
PB1884 (wt)	<u>23.80</u>	<u>39.4</u>	50.80	<u>13.9</u>	<u>23.3</u>	57.2
PB1889 (<i>SinR</i> ⁻)	<u>55.20</u>	<u>109.5</u>	154.20	<u>22.5</u>	<u>43.5</u>	56.8
PB1891 (<i>AbrB</i> ⁻)	<u>26.70</u>	<u>45.4</u>	61.80	<u>27.3</u>	<u>39.5</u>	78.4
PB1888 (<i>Npr</i> ⁻)	64.40	76.7	97.20	<u>23.3</u>	<u>31.3</u>	42.4
PB1885 (<i>Npr</i> ⁻ <i>Hpr</i> ⁻)	<u>41.40</u>	<u>60.8</u>	86.30	<u>16.4</u>	<u>28.4</u>	72.5
PB1886 (<i>Npr</i> ⁻ <i>DegU</i> ^{hy})	<u>56.80</u>	<u>63.6</u>	136.00	<u>22.2</u>	<u>30.5</u>	18.4
PB1887 (<i>Npr</i> ⁻ <i>DegQ</i> ^{hy})	83.60	125.3	162.30	<u>23.2</u>	<u>33.8</u>	35.4

* The relevant phenotype is in parentheses.

Fig. 4. Mapping of the 5' end of the *aprX* mRNA by primer extension analysis. RNA was isolated from cells of wild-type strain PB1768 grown in 2 × NB medium up to t_2 . The transcription start site is marked with an arrow. Lanes G, A, T and C show the dideoxy sequencing ladder obtained with the same primer used for primer extension: note that the gel was exposed to Amersham Hyperfilm MP film for 6 d with intensifying screens.

appropriately digested and hybridized to the integrative vector probe pJM783/455EB.

The integration of the two plasmids pJM783/232EB and pJM783/455EB placed the entire *aprX* promoter region upstream of the *lacZ* gene, and generated two different constructs downstream of the fusion, in front of the intact *aprX* gene. In the case of strain PB1883, *aprX* was preceded by the P_s region of 232 bp; in the case of strain PB1884, *aprX* was preceded by the P_L and P_s sequences (Fig. 1). Only strain PB1884 gave significant levels of β -galactosidase expression (Table 2) and enabled the analysis of *aprX* transcription. As shown in Fig. 5, *aprX-lacZ* fusion is expressed in the late-stationary

phase in both 2 × NB rich medium and TM minimal medium.

The lack of activity displayed by strain PB1883 may be due to the inability of the 232 bp fragment to promote the expression of *aprX* itself and of the downstream sequences, if an autoregulative effect is invoked. A regulatory role of AprX or of the product of the downstream gene *ymaE* could explain why in the absence of transcription of *aprX* and/or *ymaE* there is no β -galactosidase activity. Although *aprX* is followed by a potential hairpin terminator and by a weak σ^A promoter, the downstream gene *ymaE* could be under the control of the *aprX* promoter. In support of this interpretation, we observed that the induction of *ymaE*, obtained in strain BFS2601, in which this gene is under the control of the *spac* promoter of pMUTIN4 (Vagner *et al.*, 1998), causes a twofold increase of expression from the *aprX* promoter (Valbuzzi, 1998). In rich medium the presence of 1 mM IPTG increased the expression level from 4.5 to 9.8 Miller units (mg protein)⁻¹ at t_2 and from 6.5 to 11.5 at t_4 . Together, these observations suggest an operon organization for these two genes, and a possible regulatory role for YmaE. Anyhow, even if primer extension analysis (Fig. 4) showed that transcription starts from the P_s region, additional upstream sequences seem to be required in combination with the minimal promoter sequence to establish *aprX* transcription.

Since *aprX* was expressed in late-stationary phase, we evaluated the possible effect of mutations affecting some transition state regulators, by measuring the β -galactosidase activity, at different times after the t_0 , in various genetic backgrounds (Table 2). The presence of the *degU*^{hy}, *degQ*^{hy} and *hpr* mutations, which highly increase subtilisin expression (Ferrari *et al.*, 1993), did not cause analogous effects on *aprX* expression (strains PB1886, PB1887 and PB1885, compared to parental strain PB1888). The inactivation of *sinR* and *abrB* did

Fig. 5. β -Galactosidase activity of the *aprX-lacZ* transcriptional fusion. ●, ○, Enzyme activity of strain PB1884 grown, respectively, in 2 $\times$ NB and TM medium; ■, □, enzyme activity of strain PB1889 grown, respectively, in 2 $\times$ NB and TM medium. Plotted values are the mean of values obtained in two different experiments.

not stimulate the expression of *aprX* (strains PB1889 and PB1891 compared to parental strain PB1884 in Table 2 and Fig. 5). Nevertheless a significant increase of activity was observed in TM minimal medium for the *sinR* mutant (Fig. 5 and Table 2). It is not clear, however, if the negative role of SinR is exerted directly or indirectly.

An involvement of *aprX* in the sporulation process was suggested by the oligosporogenous phenotype associated with its inactivation, performed in strain BFS2601 according to the procedure for the systematic analysis of unknown genes (Vagner *et al.*, 1998; ValbuZZi, 1998). Therefore we decided to analyse in detail this aspect. *aprX* disruption was obtained in strain PB1906 by means of a more reliable construct, deriving from the insertion of the pJM105A CAT cassette by double crossover. No difference was detected in sporulation assays between parental strain 168 and PB1906 (sporulation frequency around 45% at t_{22}).

Conclusions

We have identified the *aprX* gene of *B. subtilis* strain 168, encoding a subtilisin-like serine protease. Interestingly, its primary sequence presents several Cys residues (nine over the entire protein, four in the catalytic domain). This amino acid, usually absent in subtilases from Gram-positive bacteria, may play an important role in contributing to the overall stability of the protein, via disulfide bridge formation. The absence of a canonical signal sequence in the deduced N-terminus suggests an intracellular location; the N-terminal extension, relative to the catalytic domain, with no similarity to known proteins, could play a role as a pro-peptide or as a regulator domain.

The data obtained from RT-PCR and primer extension experiments suggest that P_s region sequences are responsible for promoting transcription and, together with the data obtained by means of pJM783 transcriptional fusions, indicate that *aprX* is expressed in stationary phase. The moderate σ^A consensus displayed by the promoter sequence associated with the transcription start site could be the cause of the low *aprX* expression level, which was monitored by means of β -galactosidase assays. This could in turn explain why a long exposure time was necessary to detect the primer extension product.

The sequences upstream of *aprX* tested in pJM115 transcriptional fusions do not promote transcription of *lacZ* in the *amyE* locus ('*in trans*' position). The reason for this can be found in the fact that the putative promoter sequences assayed '*in trans*' do not comprise the entire regulatory region and/or by the fact that the promoter is not active in the ectopic position.

aprX is expressed during the stationary phase, but according to a scheme which can not be superimposed on that described for *aprE* (Pero & Sloma, 1993; Ferrari *et al.*, 1993). Only *sinR* exerts a negative effect on *aprX* transcription. The regulation of *aprX* expression resembles the control of a late-stationary-phase phenomenon. However, *aprX* is not essential for either growth or sporulation.

ACKNOWLEDGEMENTS

This work was supported by the European Commission under the Biotechnology program (BIO4-CT95-0278) and by the Università di Pavia, FAR '96 & '97.

REFERENCES

- Albertini, A. M. & Galizzi, A. (1985). Amplification of a chromosomal region in *Bacillus subtilis*. *J Bacteriol* **162**, 1203–1211.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- Anagnostopoulos, C. & Spizizen, J. (1961). Requirements for transformation of *Bacillus subtilis*. *J Bacteriol* **81**, 741–746.
- Devine, K. M. & Noone, D. (1998). Regulation of expression of the serine protease *htrA* in *Bacillus subtilis*. In *Abstracts of the International Conference on Bacilli*, p. 69. Edited by H. Yoshikawa, H. Saitou, Y. Kobayashi & I. Imamura. Senri-Chuo, Osaka, Japan.
- Ferrari, E., Jarnagin, A. S. & Schmidt, B. F. (1993). Commercial production of extracellular enzymes. In *Bacillus subtilis and Other Gram-positive Bacteria*, pp. 917–937. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.
- Glaser, P., Kunst, F., Arnaud, M. & 14 other authors (1993). *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region from 325 degrees to 333 degrees. *Mol Microbiol* **10**, 371–384.
- Harwood, C. R. (1992). *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. *Trends Biotechnol* **10**, 247–256.

- Koide, Y., Nakamura, A., Uozumi, T. & Beppu, T. (1986). Cloning and sequencing of the major intracellular serine protease gene of *Bacillus subtilis*. *J Bacteriol* **167**, 110–116.
- Kunst, F., Ogasawara, N., Moszer, I. & 148 other authors (1997). The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249–256.
- Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol* **183**, 63–98.
- Perego, M. (1993). Integrational vectors for genetic manipulations in *Bacillus subtilis*. In *Bacillus subtilis and Other Gram-positive Bacteria*, pp. 615–624. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.
- Pero, J. & Sloma, A. (1993). Proteases. In *Bacillus subtilis and Other Gram-positive Bacteria*, pp. 939–952. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.
- Perrière, G. & Gouy, M. (1996). WWW-Query: an on-line retrieval system for biological sequence banks. *Biochimie* **78**, 364–369.
- Power, S. D., Adams, R. M. & Wells, J. A. (1986). Secretion and autoproteolytic maturation of subtilisin. *Proc Natl Acad Sci USA* **83**, 3096–3100.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Schaeffer, P., Millet, J. & Aubert, J. (1965). Catabolite repression of bacterial sporulation. *Proc Natl Acad Sci USA* **54**, 704–711.
- Scotti, C., Valbuzzi, A., Perego, M., Galizzi, A. & Albertini, A. M. (1996). The *Bacillus subtilis* genes for ribonucleotide reductase are similar to the genes for the second class I NrdE/NrdF enzymes of *Enterobacteriaceae*. *Microbiology* **142**, 2995–3004.
- Shinde, U. P., Liu, J. J. & Inouye, M. (1997). Protein memory through altered folding mediated by intramolecular chaperones. *Nature* **389**, 520–522.
- Siezen, R. J. & Leunissen, J. A. M. (1997). Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Sci* **6**, 501–523.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.
- Vagner, V., Dervyn, E. & Ehrlich, S. D. (1998). A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* **144**, 3097–3104.
- Valbuzzi, A. (1998). *Analisi genetica e funzionale dei geni nrdEF (RRasI) ed aprX (subtilasi) di Bacillus subtilis nella regione compresa tra 159 ° e 180 ° del genoma*. PhD thesis. Università degli Studi di Pavia.
- Wells, J. A. & Estell, D. A. (1988). Subtilisin – an enzyme designed to be engineered. *Trends Biochem Sci* **13**, 291–297.

Received 5 February 1999; revised 3 June 1999; accepted 21 June 1999.