The wprA gene of Bacillus subtilis 168, expressed during exponential growth, encodes a cell-wall-associated protease

Philippe Margot and Dimitri Karamata

The nucleotide sequence of wprA, a protease-encoding gene of Bacillus subtilis 168, is reported. The gene, expressed during the exponential growth phase, belongs to a monocistronic operon. WprA is a 96 kDa polypeptide endowed with a signal peptide, as well as a propeptide. Upon processing and export, it gives rise to two previously identified cell-wall-bound proteins, CWBP23 and 52. Processing of WprA exhibits a novel feature of protein export, whereby removal of the middle part of the molecule accompanies the targeting to the cell wall of its N- and C-terminal parts, which correspond to CWBP23 and 52, respectively. Sequence analyses and enzymic assays reveal that CWBP52 is a serine protease. Growth rate, cell morphology, sporulation and motility of wprA mutants apparently do not differ from those of the parent strain.

Keywords: Bacillus subtilis, cell-wall-bound proteins, serine protease, polypeptide processing, chaperones

INTRODUCTION

The relatively thick cell wall of Gram-positive bacteria consists of highly cross-linked peptidoglycan (PG), to which anionic polymers are covalently linked in nearly all examined cases (Rogers et al., 1980). In addition, several proteins are associated with the bulk of the wall. Electrostatic forces are responsible for the attachment of Bacillus subtilis cell-wall-bound proteins (CWBPs) (Studer, 1988), while the linkage of Staphylococcus aureus protein A to PG is covalent (Schneewind et al., 1990). Frequently, proteinaceous paracrystalline S-layers are present at the outer surface of the wall (Sleytr & Messner, 1988).

The cell wall has a remarkable property: it is solid enough to resist high turgor pressure, but capable of relatively rapid surface extension (Rogers et al., 1980). The latter, inherent to cell growth, is apparently achieved by continuous rearrangements of the thick PG layer, by up-welling and lateral spreading (Pooley, 1976). These processes most likely require the activity of specific PG-degrading enzymes. However, not one of the many identified and genetically characterized PG hydrolases of B. subtilis 168 (Kuroda & Sekiguchi, 1990; Foster, 1991; Lazarevic et al., 1992; Longchamp et al., 1994; Margot et al., 1994; Sekiguchi et al., 1995; Smith & Foster, 1995) seems to be essential for cell growth. Margot et al. (1994) have stressed that the major endolytic activities, associated with LytC and LytD, an amidase and a glucosaminidase, respectively, are not required for extension and growth of the cell wall. However, there is evidence (Smith & Foster, 1995) that different hydrolases may substitute for each other and thus allow cell surface expansion in specific endolysin-deficient strains.

In a search for new, possibly more specific, PG hydrolases, we propose to characterize the CWBPs of B. subtilis 168 (Studer, 1988). We report here the identification of gene wprA, whose product, a 96 kDa polypeptide, is processed into two proteins, CWBP52, endowed with a protease activity, and CWBP23, the possible function of which is discussed.

METHODS

Bacterial strains, plasmids and growth. Escherichia coli strain DH5α, host for subcloning in vector pMTL20EC (Oultram et al., 1988), was grown in LB medium. Bacillus subtilis strains (Table 1) were grown in LB or MTOa medium. The latter consisted of 50 mM potassium phosphate, pH 7.0, 0.5% glucose, 0.5% Casamino acids and trace elements (Pollock, 1965), modified according to Schlegel et al. (1982). When required, media were supplemented with either 50 μg ampicillin ml⁻¹ or 3 μg chloramphenicol ml⁻¹. The sporulation medium was NB, the complex medium of Schaeffer et al. (1965). Cultures, aerated
P. MARGOT and D. KARAMATA

**Table 1. B. subtilis strains**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source or construction*</th>
</tr>
</thead>
<tbody>
<tr>
<td>168</td>
<td>trpC2</td>
<td>C. Anagnostopoulos</td>
</tr>
<tr>
<td>ATCC 6051</td>
<td>Wild-type, Marburg strain</td>
<td>ATCC collection</td>
</tr>
<tr>
<td>L16601</td>
<td>Wild-type</td>
<td>ATCC 6051 → 168 trpC2</td>
</tr>
<tr>
<td>L16623</td>
<td>wpr::Cm</td>
<td>pPM137 → L16601</td>
</tr>
</tbody>
</table>

* → Indicates transformation from donor to recipient.

by bubbling or shaking, were grown at 37 °C. The cell concentration was followed by a EEL nephelometer (Diffusion Systems). For *B. subtilis* in LB medium, a nephelometric density (ND) of 100 corresponds to 10^6 cells ml^-1.

**N-terminal amino acid sequence determination.** CWBPs were extracted from native cell wall preparations of strain L16601 with the SDS-PAGE loading buffer, separated by SDS-PAGE by bubbling or shaking, were grown at 37 °C. The cell concentration was followed by a EEL nephelometer (Diffusion Systems). DNA samples were transferred by alkali blotting (Reed et al., 1984) using a Hybond-N nylon membrane (Amersham) and labelled oligonucleotides (Amersham) as described by Monod et al., 1985) from agarose to a Hybond-N nylon membrane (Amersham), were probed with random priming method (Feinberg & Vogelstein, 1983) using [32P]-[a-35S]deoxythioadenosine triphosphate (> 1000 Ci mmol^-1, > 37 TBq mmol^-1) from Amersham. DNA from a recombinant λ was prepared with the Qiagen λ mini kit.

For Southern hybridization (Southern, 1975), size-fractionated DNA samples were transferred to alkali blotting (Reed & Mann, 1985) from agarose to a Hybond-N nylon membrane (Amersham). Probes were prepared by the random priming labelling method (Feinberg & Vogelstein, 1983) using [a-32P]dATP (3000 Ci mmol^-1; 111 TBq mmol^-1) from Amersham.

Restriction endonucleases, T4 DNA ligase, T4 DNA polymerase, calf intestinal alkaline phosphatase and Klenow polymerase were obtained from Biofinex and used according to the supplier’s instructions. DNA fragments for subcloning or for preparation of probes were recovered from agarose gels with the QIAquick gel extraction kit (Qiagen). Plasmid DNA was isolated from *E. coli* with the QIAprep spin plasmid miniprep kit. Chromosomal preparations of *B. subtilis* DNA were obtained with the Qiagen genomic kit.

**DNA sequencing.** DNA sequencing of both strands was performed by the dideoxy chain-termination method (Sanger et al., 1977) with the T7 sequencing kit (Pharmacia) according to the supplier’s instructions. 5’-[a-35S]deoxythiodenosine triphosphate (> 1000 Ci mmol^-1, > 37 TBq mmol^-1) was from Amersham. M13 primer 17mer (–20), M13 reverse primer 16mer (–24), and internal oligonucleotides were synthesized with an SM oligonucleotide synthesizer (Beckman). The sequences were analysed by the University of Wisconsin Computer Group Software (Devereux et al., 1984) and the BLAST program (Altschul et al., 1990).

**Primer extension.** RNA was prepared with the total RNA kit from Qiagen and the primer extension experiment was performed as described by Lazarevic et al., 1992.

**Transformation.** Plasmids were introduced into *E. coli* DH5α by the procedure of Chung et al., 1989. Transformation of *B. subtilis* was according to Karamata & Gross, 1970.

**Preparation of native cell walls.** Cells grown in MTOa medium at 37 °C to a concentration of 10^6 cells ml^-1 were collected by centrifugation, washed with cold double-distilled water, and disrupted by a French press (SLM Instruments). Further purification by differential centrifugation was according to Studer & Karamata, 1988. To separate CWBPs by SDS-PAGE, 1 mg lyophilized cell wall was directly resuspended in the sample buffer (Laemmli, 1970).

**Salt extraction of CWBPs.** Cells, grown and washed as described above, were extracted with 3 M NaCl in 50 mM Tris/HCl pH 8.0. After precipitation with 3 vols ethanol at –20 °C, samples of salt-extracted proteins were applied to a 12% (w/v) polyacrylamide gel in a vertical electrophoresis apparatus (Laemmli, 1970).

**Protease assay.** Cells of strains L16601 and L16623 wpr::Cm were grown in 1 litre cultures of MTOa medium. At a concentration of 10^8 cells ml^-1, bacteria were harvested and extracted with 4 ml 3 M NaCl in 50 mM Tris/HCl pH 8.0. The 600 μl reaction mixture consisted of protease buffer (50 mM Tris/HCl pH 8.0, 10 mM MgCl₂), 1 mg azocasein (Sigma), and 30 μl of the cell-wall-protein preparation. After 2 h incubation at 37 °C, the reaction was stopped by addition of 60 μl TCA (100%, w/v). After 30 min on ice, the mixture was centrifuged and 500 μl of the supernatant was used to determine the absorbance at 360 nm.

**RESULTS**

**Isolation of DNA clones encoding CWBP23 and CWBP52**

To determine their N-terminal amino acid sequence, CWBPs present in a native cell wall preparation were separated by SDS-PAGE, transferred onto a membrane,
**Fig. 2.** Nucleotide sequence of the wprA region. The deduced amino acid sequence is shown above the non-coding strand. The asterisk denotes the stop codon. The wprA putative ribosome-binding site is underlined. The vertical arrows point to the transcription start signals determined by extension of primers (see Fig. 4). >>> <<< indicates a potential terminator. The amino acid sequences of the N-termini of proteins CWBP23 and 52, obtained on purified protein preparations, are underlined.
Fig. 3. Comparison of the sequences of six serine proteases identified within the genus Bacillus. WPR, WprA (this work); SBLI, subtilisin from B. licheniformis (Jacobs et al., 1985); SBSU, subtilisin from B. subtilis (Stahl & Ferrari, 1984); THERM, thermolysin from Thermaactinomyces vulgaris (Meloun et al., 1985); ISP, Isp-1, the major internal protease from B. subtilis (Koide et al., 1986); EPR, a minor exoprotease from B. subtilis (Sloma et al., 1988). Residues common to at least four proteins are enclosed in boxes. Asp (D), His (H), Ser (S) residues belonging to the active site of subtilisins are marked with an asterisk.

and sequenced. Oligonucleotide probes were designed on the basis of these sequences and the codon usage for genes encoding the four already sequenced CWBPs (Lazarevic et al., 1992; Foster, 1993; Margot et al., 1994). Those were used to screen a B. subtilis AEMBL3 library (Mauel et al., 1989). Probes A and B, corresponding to N-terminal peptides ANDIQYPYQWP and KEQTGAMKEP of CWBP23 and 52, respectively, hybridized to the same clones. Analysis by restriction enzymes of the A clone K55 revealed that probe A hybridizes with a 6.0 kb SaGI and a 0.12 kb HindIII fragment. Subcloning of these fragments was unsuccessful, while sequencing of the 2.2 kb at the right-hand end of K55 revealed a truncated ORF. Plasmid rescue by digestion of strain L16601::pPM120 chromosomal DNA with PstI yielded plasmid pPM129 (Fig. 1) which contains the C-terminal of the ORF. The sequence upstream of pPM137 (Fig. 1) was obtained by direct sequencing of K55. To localize the cloned region, plasmid pPM120 (Fig. 1) was hybridized to the YAC library of Serrot et al. (1993). Hybridization to YAC clones 10–208 and 10–28 (Serrot et al., 1993) places the CWBP23 and 52 encoding gene(s) at around 97° on the B. subtilis genetic map, close to the add marker (Anagnostopoulos et al., 1993).

**Sequence analysis**

Nucleotide sequencing of the relevant region (Fig. 2) revealed a single ORF of 894 amino acids, preceded by a potential ribosome-binding site (ΔG = −18.6 kcal mol⁻¹; −77.8 kJ mol⁻¹) and followed by a strong hairpin loop that may act as a p-independent terminator (ΔG = −24.6 kcal mol⁻¹; −102.9 kJ mol⁻¹). The first 31
amino acids correspond to a signal peptide. A positive N-terminal region is followed by a hydrophobic core of 12 amino acids, which, in turn, is followed by a more polar region and a putative signal peptidase cleavage site, in agreement with the -1, -3 rule of von Heijne (1986). The subsequent 25 amino acids exactly match the CWBP23 N-terminal sequence, confirming the presence of the upstream signal peptide. Amino acids 414-428 match the CWBP52 N-terminal sequence. The molecular mass of the deduced polypeptide extending between amino acids 414 and the end of the polypeptide is 52 kDa, in good agreement with its apparent molecular mass of 55 kDa determined by SDS-PAGE (Studer, 1988). From now on, this protein, previously described as CWBP55 (Studer & Karamata, 1988), will be denoted CWBP52. CWBP23 and 52 appear to be produced by the cleavage of a large 96 kDa polypeptide, possibly with the excision of an intercalating polypeptide of about 20–25 kDa. The calculated isoelectric point of the 96 kDa polypeptide is 10.1. However, from the sequence, it can be deduced that its N-terminal region, corresponding to CWBP23, is acidic (pI 4.7), while the C-terminal moiety, encompassing CWBP52, is basic (pI 9.9).

Amino acid sequence homology analyses with the BLAST (Altschul et al., 1990) or FAST (Devereux et al., 1984) programs strongly suggested that CWBP52 is a serine protease of the subtilisin family (Fig. 3). Therefore, the gene encoding CWBPs 23 and 52 was tentatively named \textit{wprA} (wall protease).

Primer extension analysis revealed two start signals, located 15 and 49 bp upstream of the ribosome-binding site (Fig 4). Inspection of the sequence at positions -10 and -35, upstream of these start signals, revealed poor promoter consensus sequences. Only the TATGAT sequence, present 12 bp upstream of the most intense signal, bears homology with the canonical -10 P\textsubscript{A} sequence, 5 out of 6 residues being conserved. However, the possibility of processed mRNA (Condon et al., 1996), transcribed from promoters located further upstream, remains open.

**Phenotype of \textit{wprA}-deficient strains**

To obtain \textit{wprA}-deficient mutants, Campbell-type insertions of each of the plasmids pPM137, pPM121, pPM128 and pPM120 (Fig. 1) were performed. Salt (3 M NaCl) extracts of whole cells of the strains so obtained were devoid of both CWBP23 and 52 (Fig. 5a). Absence of CWBP23 in both the wall and the supernatant of all examined strains, even those with insertions affecting only the downstream CWBP52 protein, which should have allowed the synthesis of the proximal CWBP23, suggested
Table 2. Protease assay

<table>
<thead>
<tr>
<th>Strain</th>
<th>Protease activity (U)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No inhibitor†</td>
</tr>
<tr>
<td>L16601 wprA+</td>
<td>15.6, 14.5</td>
</tr>
<tr>
<td>L16623 wprA::Cm</td>
<td>1.4, 1.9</td>
</tr>
</tbody>
</table>

*One unit (U) corresponds to an increase in A\textsubscript{600} of 0.001 after 2 h incubation at 37 °C. The results of duplicate assays are shown.
†The reaction buffer was 50 mM Tris/HCl pH 8.0, 10 mM MgCl\textsubscript{2}, 150 mM NaCl and 1 mg Azocasein in a 600 μl reaction mixture.
‡2 mM PMSF was added to the reaction buffer.

that the intact 96 kDa protein is required for a correct processing leading to both CWBP23 and 52. The strain bearing the pPM129 insertion and the wild-type exhibited an identical CWBP pattern, confirming that the gene encoding CWBP23 and 52 ends upstream of the PstI site (Fig. 1). Electrophoresis of proteins extracted by SDS from native cell walls of strain L16623 wpr::Cm confirmed that inactivation of wprA was accompanied by the absence of both CWBP23 and 52 (Fig. 5b). So far, no other phenotype corresponding to a wprA-deficient strain has been identified. Sporulation, motility and growth rate in the media used did not differ from those of the parent strain.

Measurement of the protease activity in salt extracts

Protease assays of salt extracts from cells of the exponentially growing wild-type strain revealed an activity which was strongly reduced in the mutant lacking WprA (Table 2). This activity was almost completely inhibited by 2 mM PMSF, revealing that wprA encodes a serine protease, apparently constitutively expressed during the exponential growth phase. While the sequence analysis of the CWBP52 structural gene is in perfect agreement with the enzymic assay, the role of CWBP23 remains to be elucidated. The possibility that the latter protein is responsible for a PMSF-resistant protease activity (Table 2) is not supported by sequence analysis (see Discussion).

DISCUSSION

Growth of Gram-positive bacteria is intrinsically coupled to the reorganization of their thick cell wall (Pooley, 1976), a process probably requiring different, specific, PG hydrolases (Roten et al., 1994). To uncover such enzymes, we have undertaken a systematic study of the CWBPs of *B. subtilis* 168. The experiments reported here have led to the identification of *wprA*, a 2682 bp monocistronic operon, localized at 97° on the *B. subtilis* genome. *wprA* was shown to encode a 96 kDa polypeptide which, upon processing, gives rise to two CWBPs, i.e. 23 and 52. Analysis of sequence homologies revealed that CWBP52 is a serine protease, an observation confirmed by the loss of cell wall associated protease activity in *wprA* mutants. Comparison of well-characterized *B. subtilis* exoproteases with CWBP52, a protease which remains attached to the cell wall, revealed a novel form of processing involved in protein secretion. Observations on protease export have led to a coherent picture (Pero & Sloma, 1993). Proteases were shown to be endowed with a pre- and a propeptide. The former corresponds to a signal peptide, required by many bacterial proteins for the translocation across the cytoplasmic membrane (Gierash, 1989), whereas the latter, representing generally the N-terminal moiety, is likely to be a chaperone-type molecule cleaved during the late stages of secretion (Pero & Sloma, 1993). Propeptides seem to be a characteristic element of this class of Gram-positive exoenzymes. Although no motif specific to such elements has been identified, they apparently have in common a relatively high proportion of charged amino acids (Ohta et al., 1991). WprA seems to have the main attributes of an exoprotease: (i) its 31 N-terminal amino acids correspond to a consensus signal peptide allowing passage across the cytoplasmic membrane, and (ii) its N-terminal moiety, rich in charged amino acids [among 383 residues, 64 are positive (KR) and 53 negative (DE)], is most likely to represent a propeptide. However, the cleavage of WprA is apparently complex, since both its N- and C-terminal peptides, corresponding to previously identified CWBP23 and 52, respectively, are targeted to the cell wall. This type of processing, equivalent to the removal of the middle portion of an exoprotein, is a novel feature.

CWBP23 and 52 are attached by electrostatic forces to the cell wall, from which they can be extracted with 3 M NaCl. While CWBP52, with a calculated isoelectric point of 9.9, can easily interact with the negatively charged cell wall, the attachment of CWBP23, a negatively charged peptide (pI 4.1), cannot be readily accounted for. When a preparation of salt-extracted and desalted CWBPs is incubated with deproteinized wall, both 23 and 52 reassociate with the latter (Studer, 1988). Interestingly, autolysis of native cell walls leads to rather particular forms of CWBP23 and 52, which, unlike the other inventoried CWBPs, are unable to bind to deproteinized cell walls (Studer, 1988). This could be accounted for by assuming that, upon hydrolysis, CWBP23 and 52 remain associated with low-molecular-mass cell wall degradation products, which prevent their attachment to cell wall and possibly ensure their stability. Presently, we hypothesize that the proteins are jointly targeted to the cell wall, where they remain intimately coupled. Such an association, allowing indirect linkage of the negatively charged CWBP23 to the wall, seems to be compatible with two further observations. First, insertionally inactivation of the CWBP52, encoded by the distal part of *wprA*, is accompanied by the disappearance of CWBP52, as well as of 23, suggesting the requirement of the entire WprA for the processing and proper targeting of CWBP23. Second, sequence homology analyses, revealing beyond doubt that CWBP52 is a serine protease, suggest that CWBP23 may share a weak homology with eukaryotic protease modulators (not presented).

The biological role of CWBP23 and 52 remains open to speculation. Their location at the cell surface level
suggests a possible interaction with the surrounding medium, allowing, for instance, an immediate degradation of proteinaceous nutrients. Since wprA is expressed throughout the exponential growth phase, a phenomenon most unusual for proteases, whose synthesis is generally induced in the stationary phase, this – hypothetical – exoprotein degradation could take place without the delay required for the induction of the battery of other host proteases. Proteases present at the surface of *Streptococcus lactis* cells (Kiwaki et al., 1989) may fulfill the same role, although their association with the cell surface seems to be due to an anchoring to the cytoplasmic membrane rather than the electrostatic interaction with the PG. Finally, CWBP52 belongs to a class of rather large proteolytic enzymes, and its possible role in PG degradation, by cleavage of the peptide bridge, remains to be investigated.

**ACKNOWLEDGEMENTS**

The expertise of Dr Arthur Moir (University of Sheffield, UK), who performed the protein sequencing, and the help of the Fonds National Suisse de la Recherche Scientifique (grant no. 31-42522.94) are gratefully acknowledged.

**REFERENCES**


Received 5 June 1996; revised 24 July 1996; accepted 15 August 1996.