Gene Encoding a Novel Extracellular Metalloprotease in *Bacillus subtilis*

ALAN SLOMA,* CATHY F. RUDOLPH, GERALD A. RUFO, JR., BARBARA J. SULLIVAN, KELLY A. THERIAULT, DELPHINE ALLY, AND JANICE PERO

BioTechnica International, Inc., 85 Bolton Street, Cambridge, Massachusetts 02140

Received 24 August 1989/Accepted 20 October 1989

The gene for a novel extracellular metalloprotease was cloned, and its nucleotide sequence was determined. The gene (*mpr*) encodes a primary product of 313 amino acids that has little similarity to other known *Bacillus* proteases. The amino acid sequence of the mature protease was preceded by a signal sequence of approximately 34 amino acids and a pro sequence of 58 amino acids. Four cysteine residues were found in the deduced amino acid sequence of the mature protein, indicating the possible presence of disulfide bonds. The *mpr* gene mapped in the *cya-arol* region of the chromosome and was not required for growth or sporulation.

The gram-positive, spore-forming bacterium *Bacillus subtilis* produces and secretes proteases, esterases, and other types of exoenzymes at the end of the exponential phase of growth (15). The principal extracellular proteolytic enzymes, alkaline (subtilisin) and neutral (metallo-) proteases, are encoded by the *apr* and *npr* genes, respectively (11, 24, 26, 29). In addition, the genes for two minor extracellular proteases, *epr* (21) and *bpr* (encoding bacilcopeptidase F; A. Sloma, G. A. Rufo, Jr., C. F. Rudolph, B. J. Sullivan, K. A. Theriault, and J. Pero, submitted for publication) have been identified. Strains bearing null mutations in these four protease genes and the major intracellular protease gene *isp-I* (12) still produce extracellular protease activity. The majority of the residual protease activity can be attributed to a novel cysteine-containing metalloprotease (Mpr) that has been purified to apparent homogeneity (17). Here we report the cloning of the gene encoding this cysteine-containing protease.

MATERIALS AND METHODS

**Bacterial strains and plasmids.** *B. subtilis* strains are listed in Table 1. Plasmid pBl81/6 is a derivative of pBD64 (8). Plasmids pBR322 and pUC19 were used for cloning into *Escherichia coli* DH5 cells obtained from Bethesda Research Laboratories, Inc., Gaithersburg, Md. The cat gene was obtained from plasmid pM11101 (30), and the ble gene was obtained from pUB110 (20). *B. subtilis* strains were grown on tryptose blood agar base or minimal glucose medium and were made competent by the procedure of Anagnostopoulos and Spizizen (1). Selection for phleomycin resistance was carried out on tryptose blood agar base by the overlay method after a 1.5-h delay at 37°C to allow for expression. The final concentration of phleomycin was 2 µg/ml. Plasmid DNA from *B. subtilis* and *E. coli* was prepared by the alkaline lysis method of Birnboim and Doly (3). Plasmid DNA transformation in *B. subtilis* was performed as described by Gryczan et al. (7).

**Enzymes and chemicals.** Restriction enzymes, T4 DNA ligase, T4 polynucleotide kinase, and calf intestine alkaline phosphatase were obtained from Boehringer Mannheim Biochemicals, Indianapolis, Ind. The nick translation kit was obtained from Amersham Corp., Arlington Heights, Ill. Phleomycin was purchased from Calya S.A.R.L., Toulouse Cédex, France. Nucleotide triphosphates labeled with 32P were obtained from Du Pont, NEN Research Products, Boston, Mass.

**Amino acid sequence determination.** The N-terminal amino acid sequence of purified Mpr was determined by automated Edman degradation with subsequent identification and quantification of phenylthiohydantoin-labeled amino acids by reverse-phase high-performance liquid chromatography. Additional amino acid sequences were obtained from internal fragments of Mpr generated by trypsin digestion. The purified enzyme was incubated with trypsin, and the resulting peptides were separated and collected by using reverse-phase high-performance liquid chromatography.

**Oligonucleotide probe preparation.** Synthetic oligonucleotides were provided by the DNA Chemistry Department of BioTechnica and were synthesized by the phosphoramidite method (2) by using an Applied Biosystems 380A synthesizer. The oligonucleotides were end labeled with [γ-32P]ATP and T4 polynucleotide kinase.

**Southern blots and colony hybridization.** Southern blots (23) and colony hybridization (6) were performed as previously described. Semistringent conditions were used with all oligonucleotide probes. Nitrocellulose filters were prehybridized in 5 × SSC (1 × SSC is 0.15 M NaCl and 0.015 M sodium citrate)–1 × Denhardt solution (0.02% each of Ficoll [Pharmacia Fine Chemicals, Piscataway, N.J.], bovine serum albumin, and polyvinylpyrrolidone)–10% formamide–denatured salmon sperm DNA (100 µg/ml) at 37°C for 6 h. Hybridizations were performed by using the same solution with the addition of 32P-labeled probe (5 × 105 cpm/ml). Hybridizations were done at 37°C for 16 h, and the filters were washed with 2 × SSC-0.1% sodium dodecyl sulfate at room temperature for 30 min and then at 42°C for 1 h.

**DNA isolation and gene library construction.** Isolation of *B. subtilis* DNA was done as described previously (5). Total DNA from *B. subtilis* GP241 was digested with *Pst* I, and 3- to 4-kilobase (kb) fragments were electrophoresed from a 0.8% agarose gel and ligated to *Pst*I-digested pBR322 that had been treated with calf intestine alkaline phosphatase (1 U) for 30 min at 37°C and then 30 min at 50°C. The ligation was done at a ratio of insert DNA to vector DNA of 4:1. The ligation mixture was incubated at 16°C overnight and transformed into *E. coli* DH5. Approximately 10,000 colonies

* Corresponding author.
resulted, and plasmid screening indicated that 60% of the colonies contained plasmids with 3- to 4-kb inserts.

Restriction fragments to be sequenced were ligated into the appropriate sites of M13 vectors. DNA sequencing was performed by the dideoxy-chain termination method (18).

Mapping of the mpr gene. Mapping of the mpr locus was performed by PBS1 transduction (10) with a lysate from B. subtilis GP261. Cm\(^+\) transductants were scored for linkage to several loci from the set of reference strains described by Dedonder et al. (4).

### TABLE 1. Bacterial strains

<table>
<thead>
<tr>
<th>B. subtilis strain</th>
<th>Genotype and relevant characteristics</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP216</td>
<td>Δapr Δnpr Δisp-I Δepr amyE met</td>
<td>(21)</td>
</tr>
<tr>
<td>GP238</td>
<td>Δapr Δnpr Δisp-I Δepr cat</td>
<td>Sloma et al., submitted</td>
</tr>
<tr>
<td>GP239</td>
<td>Δapr Δnpr Δisp-I Δepr amyE met</td>
<td>Sloma et al., submitted</td>
</tr>
<tr>
<td>GP241</td>
<td>Δapr Δnpr Δisp-I Δepr Δhpŕ amyE met</td>
<td>Sloma et al., submitted</td>
</tr>
<tr>
<td>GP263</td>
<td>Δapr Δnpr Δisp-I Δepr Δhpŕ Δmpr amyE met</td>
<td>This work</td>
</tr>
<tr>
<td>GP264</td>
<td>Δapr Δnpr Δisp-I Δepr Δhpŕ Δmpr [sacQ*] amyE met</td>
<td>This work</td>
</tr>
</tbody>
</table>

\(^a\) hpř is a hyperproducing protease mutation (9).

Protease activity measurements. Strains were grown in modified Lactobacilli MRS Broth (Difco Laboratories, Detroit, Mich.), substituting 1.5% maltose as a carbon source. Protease activity was routinely determined by using azocoll (Sigma Chemical Co., St. Louis, Mo.; Calbiochem-Behring, La Jolla, Calif.) as a substrate. Tubes containing 10 mg of azocoll and appropriate volumes of 50 mM Tris-5 mM CaCl\(_2\) (pH 8.0) were preincubated for 30 min at 55°C with constant shaking. After preincubation, designated amounts of culture supernatant were added to the tubes containing substrate and buffer. The reaction was carried out at 55°C for 1 h. A control without enzyme was run with each set of reactions. At the end of incubation, tubes were centrifuged to remove unhydrolyzed azocoll, and the \(A_{520}\) of the resulting supernatant was determined. Plots of absorbance versus the amount of protein were linear up to an absorbance reading of 2.0. One unit of activity was arbitrarily designated as the amount of protein which yields an \(A_{520}\) of 0.5.

Sporulation. Liquid cultures were grown in DSM (19) for 24 h at 37°C. The cultures were then diluted, heated at 80°C for 10 min, and plated to determine the number of heat-resistant spores.

**RESULTS**

Oligonucleotide hybridization probes for the metalloprotease gene. To clone the mpr gene, our initial strategy was to synthesize a specific oligonucleotide probe ("guess-mer") on the basis of the determined N-terminal amino acid sequence of the purified protein. Automated sequential Edman degradation of purified Mpr (17) revealed the following N-terminal amino acid sequence: Ser-Ile-Ile-Gly-Thr-Asp-

![FIG. 1. Determined amino acid sequences of tryptic fragments of Mpr and corresponding oligonucleotide guess-mers. Indicated are separate amino acid sequences of three fragments. The corresponding nucleotide sequences of the synthesized oligonucleotides (guess-mers) used as DNA probes are indicated.](http://jb.asm.org/)
Bacillus resulting codons. degenerate three amino acid sequence Glu-Arg-Thr-Arg-Ile-Ser-Ser-Thr-Thr.

Since this sequence contained three serine and two arginine residues, any probe would be based on a sequence with a high degree of degenerate codons. To obtain additional amino acid sequence that could be used to construct a more specific probe for mpr, the protease was digested with trypsin, and three of the resulting internal fragments (T90, T94, T92) were isolated and sequenced (Fig. 1). DNA oligonucleotide guesses were designed and synthesized on the basis of the amino acid sequences and by relying on codon usage in Bacillus spp. for nucleotides that were uncertain because of the degeneracy of the codons. Three probes were synthesized: a 45-mer (probe BRT90) based on the amino acid sequence of T90, a 75-mer (probe 707) based on T94, and a 39-mer (probe 715) based on T92 (Fig. 1).

Probes BRT90 and 707 were labeled with [γ-32P]ATP and hybridized to duplicate Southern blots of B. subtilis GP241 chromosomal DNA digested with several restriction enzymes. The two probes hybridized to same restriction fragments: HincII, ~1 kb; PslI, 3 to 4 kb; and EcoRI, 6 to 7 kb (Fig. 2). The last probe (715) was also found to hybridize to these fragments (data not shown). Since the probes hybridized to the same restriction fragments, they were judged to be specific for the mpr gene.

Cloning of the metalloprotease gene. Since the probes

FIG. 2. Southern hybridization of the mpr gene with two guess-mer probes. Identical Southern blots of total B. subtilis chromosomal DNA digested with HincII (lane A), PslI (lane B), and EcoRI (lane C) were hybridized with labeled BRT90 probe (I) and 707 probe (II), as described in Materials and Methods.

FIG. 3. Restriction map of the 3.6-kb PslI insert of pLP1 containing mpr.

FIG. 4. Nucleotide and deduced amino acid sequences of the mpr gene. Nucleotides are numbered starting with the A of the presumed initiation codon ATG. The putative ribosome-binding site and the determined amino acid residues of the N terminus and three peptides are underlined.

Probes BRT90 and 707 were labeled with [γ-32P]ATP and hybridized to duplicate Southern blots of B. subtilis GP241 chromosomal DNA digested with several restriction enzymes. The two probes hybridized to same restriction fragments: HincII, ~1 kb; PslI, 3 to 4 kb; and EcoRI, 6 to 7 kb (Fig. 2). The last probe (715) was also found to hybridize to these fragments (data not shown). Since the probes hybridized to the same restriction fragments, they were judged to be specific for the mpr gene.

Cloning of the metalloprotease gene. Since the probes

FIG. 5. Alignment of the amino acid sequence of the Mpr protease with human protease E (HPROE) and bovine procarboxypeptidase A complex component III (PCXYA). Boxes indicate regions of homology. The numbers of the amino acid residues from each protein are shown in parentheses. Gaps were introduced to permit maximal matching.
hybridized to a 3- to 4-kb PstI fragment (Fig. 2), a gene bank was constructed of size-selected PstI fragments. To construct the bank, total chromosomal DNA from *Bacillus* strain GP241 was digested with PstI and size-fractionated fragments of 3 to 4 kb were cloned into the PstI site of pBR322. Using labeled probe 707, 1 clone out of 1,400 was identified as hybridization positive by colony and Southern hybridization analyses. This colony contained a plasmid (pLP1) with a 3.6-kb insert containing the *mpr* gene, as determined by DNA sequencing.

Characterization of the *mpr* gene. A restriction map of the 3.6-kb insert is shown in Fig. 3. Hybridization studies indicated that the three probes for *mpr* hybridized to a 1.3-kb *HincII* fragment (Fig. 2). The DNA sequence of the *HincII* fragment revealed an open reading frame that spanned most of the fragment (position -24 to position 939 in Fig. 4). The most probable initiation codon for this open reading frame is the ATG at position 1 in Fig. 4. This ATG is preceded by a *B. subtilis* ribosome-binding site (AAAGGAGG), which has a calculated ΔG of -16.0 kcal (ca. -66.9 kJ) (25). The first 33 amino acids following this Met resemble a *B. subtilis* signal sequence, with a short sequence containing four positively charged amino acids, followed by 18 hydrophobic residues, a helix-breaking proline, and a typical Ala-X-Ala signal peptidase cleavage site (14). After the presumed signal peptidase cleavage site, a propeptide of 58 residues is followed by the beginning of the mature protein as determined by the N-terminal amino acid sequence of the purified protein. The determined amino acid sequence of the peptides matched the deduced amino acid sequence except for a serine residue encoded by nucleotides 379 to 381 and a cysteine residue encoded by nucleotides 391 to 393. The determined amino acid sequences predicted a cysteine residue (position 14, T94 peptide) and an asparagine residue (position 18, T94 peptide), respectively (Fig. 1). The entire mature protein was deduced to contain 221 amino acids with a predicted molecular mass of 23,941 daltons. This size was in approximate agreement with the determined molecular mass of the purified protein of 28,000 daltons (17).

The deduced amino acid sequence showed only limited homology to other sequences in GenBank. The strongest homology was to human protease E (28) and bovine procarboxypeptidase A complex component III (27) in a 25-amino-acid sequence within Mpr (131 to 155, encoded by nucleotides 391 to 465; Fig. 4) (Fig. 5).

**TABLE 2. Determination of extracellular protease levels in supernatants of *B. subtilis* strains**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Protease (U/ml)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP238.</td>
<td>0.10</td>
</tr>
<tr>
<td>GP239.</td>
<td>80.00</td>
</tr>
<tr>
<td>GP264.</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>GP264(pCR140)</td>
<td>1.490.00</td>
</tr>
<tr>
<td>GP264(pBS810)</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

* Protease was measured by using the standard azocoll assay. One unit was defined as the amount of protein which yields an A520 of 0.5. Protease activity of supernatants was determined 20 h after the onset of stationary phase.
Deletion of mpr in the chromosome. A deletion of the mpr gene in the chromosome was constructed by replacing the wild-type gene in the chromosome with an in vitro-created deletion mutation. This deletion mutation was created by replacing the 2-kb HpaI fragment of pLP1 containing the mpr gene with a gene encoding an antibiotic resistance marker. Two different drug resistance genes were used: a cat gene was used to create pLP2 and a bleomycin resistance gene (ble) was used to construct pCR125 (Fig. 6). pLP2 and pCR125 were then linearized by digestion with EcoRI and used to transform *B. subtilis* GP241 to chloramphenicol resistance and phleomycin resistance, respectively. One chloramphenicol-resistant (GP261) and one phleomycin-resistant (GP263) transformant were selected for further study.

Cm<sup>+</sup> and Phleo<sup>+</sup> transformants were expected to be the result of a double crossover between the linear plasmid and the chromosome (marker replacement). Southern hybridization was used to confirm that the cat and ble genes had integrated in the chromosome in place of the mpr gene and that both strains carried a deletion mutation in mpr. In GP261, substitution of the cat gene for the HpaI fragment containing mpr produced a 3.2-kb *Pst*I fragment, rather than the original 3.6-kb *Pst*I fragment. In GP263, integration of the ble gene, which contained a *Pst*I site at one end, created two *Pst*I fragments of 0.5 and 2.3 kb. The expected fragments were detected when probed with nick-translated pLP1, confirming the deletion of mpr in the *B. subtilis* chromosome (Fig. 7).

Strains containing the wild-type mpr gene and those deleted for mpr were grown in liquid cultures, and the protease levels were compared. *B. subtilis* GP238 (Mpr<sup>+</sup>, mutated for five other proteases) produced low levels of protease (Table 2). The addition of sacQ<sup>+</sup> (a modified positive regulatory gene; 22), to this strain dramatically increased protease levels (*B. subtilis* GP239 [Table 2]). Most (>95%) of this protease activity was dithiothreitol sensitive (data not shown), indicating that the activity was due to Mpr (17). As expected, strains containing the deletion in mpr and sacQ<sup>+</sup> (strain GP264) produced no detectable protease activity when azocoll was used as the substrate, confirming the deletion of mpr.

No significant differences between the capacities of *B. subtilis* GP261 (Mpr<sup>+</sup>) and GP263 (Mpr mutant) to grow in MRS broth or to sporulate in Difco sporulation medium were observed, establishing that Mpr was not required for growth or sporulation.

**Location of mpr on the *B. subtilis* chromosome.** To map mpr on the *B. subtilis* chromosome, we used *B. subtilis* GP261, described above, which contained the cat gene inserted into the chromosome at the site of the mpr gene, and phage PBS1 transduction to determine the location of the cat insertion. Mapping experiments indicated that the inserted cat gene (hence, mpr) was linked to cysA14 (7% cotransduction) and to aroI906 (36% cotransduction) but unlinked to purA16 and dal. This data indicated that the mpr gene was between cysA and aroI in an area of the genetic map not previously known to contain protease genes.

**Expression of the mpr gene on a high-copy plasmid.** The mpr gene was placed on a high-copy plasmid to determine whether production of the metalloprotease would be increased. Fig. 8 shows the construction of pCR140, carrying the 2-kb HpaI fragment on a high-copy *Bacillus* plasmid, pBS81/6. *B. subtilis* GP264 carrying pCR140 produced 10- to 20-fold more protease than strains with a wild-type copy of mpr (GP239) in the chromosome (Table 2). This activity was inhibited by dithiothreitol (data not shown), indicating that the overexpressed protease was the metalloprotease.

**DISCUSSION**

*B. subtilis* strains carrying deletion mutations in four protease genes (apr, npr, isp-l, and epr) still produce extracellular protease activity (21). We have determined that this residual activity is due to at least two proteases, bacillopeptidase F (13, 16) and a novel metalloprotease of 28,000 daltons, Mpr (17).

Previously, we cloned and deleted the structural gene for bacillopeptidase F (Sloma et al., submitted). In this study, we report the cloning of the gene (mpr) for the novel metalloprotease. The mpr gene, like the gene for other extracellular proteases in *B. subtilis* (21, 24, 29), encodes a signal sequence and a propeptide that is processed during secretion. The mature enzyme of 221 amino acids corresponds well to the reported molecular mass of 28,000 daltons (17). The mature enzyme contains four cysteine residues, suggesting the possible presence of disulfide bonds. This could explain the observation that Mpr is inhibited by the reducing agent dithiothreitol (17). The presence of the cysteine residues is unique among the studied extracellular proteases of *B. subtilis* (21, 24, 29).

Mpr showed the strongest similarity to human protease E.
and bovine carboxypeptidase A complex component III (27) (Fig. 5) in a region surrounding two cysteine residues reported to participate in disulfide bond linkages in these other proteases. Surprisingly, Mpr showed no significant homology to any previously characterized bacterial proteases. It seems likely that there should exist protease genes in other gram-positive organisms that bear a similar mpr. Since this metalloprotease (Mpr) was discovered only after the removal of other major and minor proteases from Bacillus subtilis, putative homologs in other bacteria may not yet have been isolated. It will be interesting to use mpr as a probe to detect homologous proteases in other bacteria.

ACKNOWLEDGMENTS

We thank William S. Lane at the Harvard Microchemistry Department for protein sequencing. We also thank Abdul Ally for his assistance with DNA sequence analysis, Jenna Kakimoto for technical assistance, and Agnes Huot for her expert assistance in preparing the figures and manuscript. We also thank Richard Losick for helpful discussions and for assistance in preparation of the manuscript.

LITERATURE CITED