Structure and Mechanism of Action of the Protease That Degrades Small, Acid-Soluble Spore Proteins during Germination of Spores of Bacillus Species

CLAUDIO NESSI,1 MARK J. JEDRZEJAS, 2 AND PETER SETLOW1*

Department of Biochemistry, University of Connecticut Health Center, Farmington, Connecticut 06032,1 and Department of Microbiology and Center for Macromolecular Crystallography, University of Alabama at Birmingham, Birmingham, Alabama 352942

Received 27 March 1998/Accepted 22 July 1998

The germination protease (GPR) of Bacillus megaterium initiates the degradation of small, acid-soluble proteins during spore germination. Trypsin treatment of the 46-kDa GPR zymogen (termed P46) removes an ~15-kDa C-terminal domain generating a 30-kDa species (P30) which is stable against further digestion. While P30 is not active, it does autoprocess to a smaller form by cleavage of the same bond cleaved in conversion of P46 to the active 41-kDa form of GPR (P41). Trypsin treatment of P41 cleaves the same bond in the C-terminal part of the protein as is cleaved in the P46→P30 conversion. While the ~29-kDa species generated by trypsin treatment of P41 is active, it is rapidly degraded further by trypsin to small inactive fragments. These results, as well as a thermal melting temperature for P41 which is 13°C lower than that for P46 and the unfolding of P41 at significantly lower concentrations of guanidine hydrochloride than for P30, are further evidence for a difference in tertiary structure between P46 and P41, with P46 presumably having a more compact stable structure. However, circular dichroism spectroscopy revealed no significant difference in the secondary structure content of P46 and P41. The removal of ~30% of P46 or P41 without significant loss in enzyme activity localized GPR’s catalytic residues to the N-terminal two-thirds of the molecule. This finding, as well as comparison of the amino acid sequences of GPR from three different species, analysis of several site-directed GPR mutants, determination of the metal ion content of purified GPR, and lack of inhibition of P41 by a number of protease inhibitors, suggests that GPR is not a member of a previously described class of protease.

Between 10 and 20% of total spore protein is degraded to amino acids in the first minutes of germination of spores of Bacillus species (30). The proteins degraded are a group of small, acid-soluble proteins (SASP) unique to the spore stage of the life cycle. The α/β-type SASP, which are coded for by a multigene family, are bound to the dormant spore’s DNA and provide a significant component of spore resistance to heat, hydrogen peroxide, and UV radiation (28, 30, 31). The γ-type SASP, which are coded for by a single gene, are also in the spore core (the site of spore DNA) but are not bound to any macromolecule. SASP degradation at the beginning of spore germination both frees up the DNA for transcription and provides amino acids that are used for protein synthesis during subsequent development (22, 30).

SASP degradation during spore germination is initiated by one or two endopeptolytic cleavages catalyzed by a sequence specific germination protease termed GPR which is synthesized during sporulation at about the same time as its SASP substrates (30). GPR is synthesized as a protein of 46 kDa as measured by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE); this species, termed P46, is inactive (13). Two hours later in sporulation, P46 undergoes intramolecular autoprocessing which removes an N-terminal propeptide of 15 (Bacillus megaterium) or 16 (B. subtilis) amino acids; the new, fully active form of GPR is called P41 (13, 24). Both P46 and P41 are tetramers, and only the tetrameric form of P41 is active (12). The autoprocessing of P46 to P41 is triggered both in vivo and in vitro by low pH, dipicolinic acid, and dehydration (8). Although active in vitro, P41 carries out no SASP degradation during sporulation because the conditions inside the developing and dormant spore (in particular the dehydration and mineralization) are not favorable for enzymatic action. However, in the first minutes of spore germination, the spore core rehydrates, allowing rapid attack of GPR on SASP. Synthesis of GPR in an inactive form and its autoprocessing at the correct time in sporulation are essential to generate a fully resistant spore; a strain with a GPR variant that autoprocesses P46 to P41 ~1 h earlier during sporulation produces spores that are more sensitive than wild-type spores to a variety of treatments, because the α/β-type SASP content of the mutant spores is reduced (6).

In addition to the autoprocessing reaction converting P46 to P41, the P41 of B. megaterium also undergoes an additional autoprocessing to P30 with loss of an additional seven N-terminal residues (8). The enzymatic activities of P30 and P41 are identical, and P30 is not generated from P41 of B. subtilis. Consequently, the generation of P30 with B. megaterium GPR may not have any functional significance.

While the role played by GPR in SASP degradation as well as its autoprocessing are reasonably well understood, little is known of the structural differences between P46 and P41 and the nature of GPR’s active site(s). Indeed, the precise class of protease to which GPR belongs has not been established. To obtain more information on these latter topics, we have analyzed the trypsin digestion products generated from P46 and P41 and have used a variety of techniques in an attempt to determine the class of protease to which GPR belongs. The latter data suggest that GPR is not a member of a previously described class of proteases.
confirmed the presence of the desired mutations and that they were the only mutations present in the 211-bp Clai-Eegl fragment of \( B. megaterium \) gpr (Fig. 1A). This Clai-Eegl fragment was purified by agarose gel electrophoresis and ligated to Clai-Eegl-digested pPS1910 which had been purified to obtain the largest fragment. The ligation mixture was used to transform \( E. coli \) UT481 containing pPS1910 harboring an ampicillin resistance, and plasmid DNA from several clones was isolated and sequenced to ensure that gpr had acquired the desired mutation. \( E. coli \) UT481 strains carrying plasmids with the gpr S219A mutation and the gpr S229A mutation were called PS2577 and PS2578, respectively.

### MATERIALS AND METHODS

#### Bacteria and plasmids used and isolation of DNA

The bacterial strains and plasmids used in this work are listed in Table 1. \( E. coli \) strains JM83 and PS2577 were chosen for plasmid overproduction and purification.

#### Site-directed mutagenesis

We used the megaprim-PCR-based PCR method (26) to make mutations in \( B. megaterium \) gpr in which Ser\(^{274}\) or Ser\(^{275}\) is replaced by Ala (Fig. 1A). The oligonucleotide primers used for the first round of PCR were A (5'-CCATGACACTACGACCGG-3'), B (5'-CATACGGCGGTTGGGAACAA-3'), and C (5'-GCGTTAAAGAATCCTGATGAAACCTTGCG-3'). Oligonucleotide A is in the T7 promoter of the pT719U portion of pPS1907 (Table 1); oligonucleotide B begins at nucleotide (nt) 753 and ends at nt 776 and oligonucleotide C begins at nt 779 and ends at nt 809 of the coding sequence of \( B. megaterium \) gpr. Oligonucleotides B and C are complementary to the \( gpr \) sequence except for the underlined nucleotides. In the first round of PCR, the reaction mixture (100 nl) contained 100 pmol of primer B or C and 100 pmol of primer A, 1.5 mM MgSO\(_4\), 0.2 mM each deoxynucleoside triphosphate, 200 ng of template (plasmid pPS1907), and 10 \( \mu \)l of 10X Vent DNA polymerase buffer. Reaction mixtures were overlaid with mineral oil and heated at 80 \( ^\circ \)C for 5 min. At an optical density of 0.4, 20 mg/ml isopropyl-\( \beta \)-D-galactopyranoside (\( \beta \)-IPTG) was added, and samples were subjected to 35 cycles of PCR (1 min at 94 \( ^\circ \)C, 1 min at 55 \( ^\circ \)C, and 1 min at 72 \( ^\circ \)C) and incubated at 72 \( ^\circ \)C for 10 min at the end of the last cycle. The synthesized megaprimers were analyzed by electrophoresis on a 1.2% agarose gel to verify that they were of the expected sizes and then used for the second round of PCR. The second round of PCR was carried out in 50 \( \mu \)l containing 50 ng of one of the megaprimers, 500 pmol of primer D (5'-GTAAGAGACGAGCGAGCCTG-3'), which is complementary to the sequence of pT719U in pPS1907 downstream of the multiple cloning site, 1.5 mM MgCl\(_2\), 0.2 mM each deoxynucleoside triphosphate, 200 ng of HindIII-linearized pPS1907 purified by agarose gel electrophoresis, and 5 \( \mu \)l of 10X Taq DNA polymerase buffer. Samples were treated as described above, 1 \( \mu \)l of \( \beta \)-IPTG DNA polymerase was added, and PCR was performed as described above. The final PCR products were subjected to agarose gel electrophoresis; fragments with the expected sizes were excised and ligated into plasmid pRSII (Invitrogen). The ligation mixture was used to transform \( E. coli \) INVoS\(_e\) cells (Invitrogen), and transformants were selected on 2X YT agar plates containing ampicillin (50 \( \mu \)g/ml) and 5-bromo-4-chloro-3-indolyl-\( \beta \)-galactopyranoside (40 \( \mu \)g/ml). Plasmids were isolated from white colonies, and digestion with HindIII and HindIII-identi
dified plasmids with the appropriate ~1.5-kb insert. DNA sequence analysis

#### Protein overexpression and purification

\( B. megaterium \) gpr was purified as described previously (16, 17). All other strains for overexpression of GPR were grown at 37 \( ^\circ \)C in 2X YT medium plus ampicillin (50 \( \mu \)g/ml). At an optical density of 0.6-0.8, isopropyl-\( \beta \)-D-thiogalactopyranoside was added to 1 mM, and the cells were harvested after 2 h of further growth at 37 \( ^\circ \)C. GPR was routinely purified from these cells as described (6, 17).

The procedure used to purify GPR for analysis of metal ions was a modified version of the published procedure that gives a higher yield and better purity. Cells from 2 liters of culture of either PS2479 or \( E. coli \) JM83 carrying pPS1910 were harvested after 2 hours of growth at 37 \( ^\circ \)C. GPR was purified from these cells as described (6, 17).

The final PCR products were purified by agarose gel electrophoresis, and 5 \( \mu \)l of 10X Taq DNA polymerase buffer. Samples were treated as described above. The final PCR products were subjected to agarose gel electrophoresis; fragments with the expected sizes were excised and ligated into plasmid pRSII (Invitrogen). The ligation mixture was used to transform \( E. coli \) INVoS\(_e\) cells (Invitrogen), and transformants were selected on 2X YT agar plates containing ampicillin (50 \( \mu \)g/ml) and 5-bromo-4-chloro-3-indolyl-\( \beta \)-galactopyranoside (40 \( \mu \)g/ml). Plasmids were isolated from white colonies, and digestion with HindIII and HindIII-identified plasmids with the appropriate ~1.5-kb insert. DNA sequence analysis

#### Table 1. Bacterial strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant genotype or characteristics</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JM83</td>
<td>lac(^{I+})</td>
<td>This work</td>
</tr>
<tr>
<td>PS2479</td>
<td>PPS427 Amp(^{\alpha})</td>
<td>This work</td>
</tr>
<tr>
<td>PS2549</td>
<td>PPS254 Amp(^{\alpha})</td>
<td>This work</td>
</tr>
<tr>
<td>PS2569</td>
<td>PPS256 Amp(^{\alpha})</td>
<td>This work</td>
</tr>
<tr>
<td>PS2577</td>
<td>PPS2576 Amp(^{\alpha})</td>
<td>This work</td>
</tr>
<tr>
<td>PS2578</td>
<td>PPS2578 Amp(^{\alpha})</td>
<td>This work</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pPS740</td>
<td>pUC12 with ( B. megaterium ) wild-type gpr; Amp(^{\alpha})</td>
<td>32</td>
</tr>
<tr>
<td>pPS1907</td>
<td>pTIZ19 with ( B. megaterium ) gpr from which all propeptide residues are deleted; Amp(^{\alpha})</td>
<td>32</td>
</tr>
<tr>
<td>pPS1910</td>
<td>pUC12 with ( B. megaterium ) gpr from which all propeptide residues are deleted; Amp(^{\alpha})</td>
<td>6</td>
</tr>
<tr>
<td>pPS2339</td>
<td>pUC12 with ( B. megaterium ) gpr lacking propeptide residues 2 to 4; Amp(^{\alpha})</td>
<td>17</td>
</tr>
<tr>
<td>pPS2340</td>
<td>pUC12 with ( B. megaterium ) gpr lacking propeptide residues 2 to 7; Amp(^{\alpha})</td>
<td>17</td>
</tr>
<tr>
<td>pPS2341</td>
<td>pUC12 with ( B. megaterium ) gpr lacking propeptide residues 2 to 10; Amp(^{\alpha})</td>
<td>17</td>
</tr>
<tr>
<td>pPS2342</td>
<td>pUC12 with ( B. megaterium ) gpr lacking propeptide residues 2 to 13; Amp(^{\alpha})</td>
<td>17</td>
</tr>
<tr>
<td>pPS2479</td>
<td>pPS1910 with ( B. megaterium ) gpr with KDALEN instead of KDIALEN in the P(<em>{41})-T(</em>{41}) cleavage site; Amp(^{\alpha})</td>
<td>17</td>
</tr>
<tr>
<td>pPS2524</td>
<td>pPS1910 with ( B. megaterium ) gpr C115A; Amp(^{\alpha})</td>
<td>17</td>
</tr>
<tr>
<td>pPS2569</td>
<td>pPS1910 with ( B. megaterium ) gpr S274A S276A; Amp(^{\alpha})</td>
<td>17</td>
</tr>
<tr>
<td>pPS2577</td>
<td>pPS1910 with ( B. megaterium ) gpr S219A; Amp(^{\alpha})</td>
<td>17</td>
</tr>
<tr>
<td>pPS2578</td>
<td>pPS1910 with ( B. megaterium ) gpr S229A; Amp(^{\alpha})</td>
<td>17</td>
</tr>
</tbody>
</table>

\( ^{\alpha} \) Amp\(^{\alpha}\), ampicillin (50 \( \mu \)g/ml) resistant.

#### Acknowledgments

The authors thank Dr. K. deLaporte for helpful discussions. The research was supported by National Science Foundation grants DBI-9521406 and DBI-9875018.
GPR assay and autoprocessing. GPR was assayed as described previously (12). When trypsin-digested P46 or P41 was assayed, soybean trypsin inhibitor was added to digest in a 2:1 molar ratio with trypsin prior to GPR assays. When crude extracts were to be assayed, GPR was overexpressed as described above and the cells were harvested by centrifugation. The cells from ~50 ml of culture were resuspended in 2 ml of 50 mM Tris-HCl (pH 7.4)–5 mM CaCl2–20% glycerol, disrupted by sonication, and centrifuged at 12,000 g at 4°C for ~15 min. The protein concentration of the supernatant fluid was determined by the method of Bradfords (12).

Effect of protease inhibitors on GPR activity. Two different procedures were used to test the effects of protease inhibitors on the activity of purified P46. To test the effect of 1,10-orthophenanthroline (Ophen), P41 (0.2 mg/ml) was dialyzed overnight at 4°C against 50 mM Tris-HCl (pH 7.4)–5 mM CaCl2–20% glycerol–2 mM Ophen prior to enzyme assays (1). To test the effects of 250 mM N-acetylimidazole (NAI), 1 mM 3,4-dichloroisocoumarin (DCI), 10 mM disopropylfluorophosphate (DFP), 10 mM 8-hydroxyquinoline-5-sulfonic acid (HQSA), 1 mM iodoacetamide (IAA), 1 mM N-ethylmaleimide (NEM), 10 μM pepstatin, and 3 or 10 mM phenylmethylsulfonyl fluoride (PMSF), the inhibitors were routinely incubated in 50 mM Tris-HCl (pH 7.4)–5 mM CaCl2–20% glycerol with 20 to 200 μg of purified P46 per ml. Incubation was for 4 h at 23°C for NAI (the buffer used was 25 mM MOPS [pH 7.4]), 2 h at 23°C for DCI (with no glycerol in the incubation), 45 min at 37°C for DFP [50 mM 3-(cyclohexyloxyamino)-1-propanesulfonic acid (pH 10) was used with this inhibitor in addition to Tris-HCl], 30 min at 37°C for HQSA, 1 h at 23°C for IAA, overnight at 23°C for NEM, 1 h at 23°C for pepstatin (with 20 μg of P46 per ml), and 45 min at 23°C for PMSF. Aliquots of the incubation mixtures were assayed for GPR activity as described above both before and after incubation, and in parallel with P46 samples incubated similarly but without inhibitors. In some cases the discontinuous colorimetric assay for GPR (12, 29) was used, as some inhibitors inhibit the aminopeptidase used in the GPR assay.

Metal analysis. GPR for metal analysis was purified and concentrated as described above and then dialyzed exhaustively against 10 mM Tris-HCl (pH 7.4)–10% glycerol–1 mM CaCl2 in otherwise metal-free water (Milli-Q); this dialysis buffer had <2 μM Mg2+, Mn2+, Zn2+, Co2+, or Fe2+. Standard solutions were made with the dialysis buffer as the diluent, and metal analyses were carried out on a Leeman PS 1000 ICP atomic absorption spectrometer. The instrument was calibrated, and standards were checked to verify the calibration.

RESULTS

Proteolytic digestion of GPR. Previous work has indicated that the only covalent change in conversion of P46 to P41 is removal of 15 (B. megaterium) or 16 (B. subtilis) amino-
were generated (Fig. 4). Analysis of the N-terminal sequences of these two species gave MLAVE for the larger intermediate and DALAN for the smaller one (note that this experiment used ΔP41). While the first sequence is of intact P41, the second is generated by cleavage after K21 (Fig. 1B). The molecular masses of the two species as determined by MALDI-TOF were 27,302.9 and 26,760.8 Da, respectively, in fairly good agreement with the molecular masses calculated (27,309.4 and 26,566.4 Da), assuming that a second tryptic cleavage takes place after K266 as in generation of P30 from P41. While these two fragments were active (Fig. 4), their activity was lost within at most a few hours, even if samples were frozen (data not shown).

In previous studies, we showed that deletion of up to nine residues of the B. megaterium GPR propeptide did not result in a protein with significant activity, while deletion of 12 of the propeptide residues gave a protein with 30% of the activity of P41 (17). Digestion of GPR variants lacking three or six propeptide residues with trypsin gave essentially the same result as did digestion of P46; i.e., a stable inactive species of ~30 kDa was generated (data not shown). In contrast, digestion of the variants lacking 9 and 12 propeptide residues with trypsin gave results similar to those with P41, i.e., rapid digestion of the proteins with generation of short-lived active intermediates (data not shown).

**Analysis of P46 and P41 structures.** The differences in sulfhydryl reactivity and trypsin sensitivity of P46 and P41 (7) indicate that these proteins differ somewhat in structure; this variance could reflect a difference in secondary structure, in tertiary structure, or in both. We used CD spectroscopy to assess these secondary structure content of P41 and P46. The two proteins had very similar far-UV CD spectra, including characteristics found in proteins with α helices (Fig. 5). Analysis of the spectra between 190 and 260 nm by the method of Chang et al. (3) indicated that P41 contained 21% ± 5% α-helix structure, 58% ± 5% β sheet, and 21% ± 5% random coil; for P46, the corresponding values were 19% ± 5%, 51% ± 5%, and 28% ± 5%.

Although the CD spectroscopy would not have detected small differences in secondary structure between P46 and P41, the foregoing data indicate that these proteins have very similar secondary structures. This finding further suggests that the differences in trypsin sensitivity and sulfhydryl group reactivity of P41 and P46 are due to differences in the tertiary structures of these two proteins. To study this point further, we measured the thermal melting of P41 and P46 as described in Materials and Methods and found that P41 was significantly less stable than P46, their melting temperatures being 63 and 76°C, respectively (Fig. 6). The profile of the thermal unfolding of P41, as seen by CD analysis, was quite different and more irregular than that of P46, and P41 unfolding was initiated at a significantly lower temperature. Using another method to compare...
the stabilities of P₄₀ and P₄₁, we calculated that the GuHCl concentrations needed for 50% unfolding of P₄₁ and P₄₀ (determined as described in Materials and Methods) were 1.4 M for serine proteases, the four main classes of proteases (2, 19–

FIG. 5. CD spectra of P₄₁ and P₄₆. The far-UV spectra of P₄₆ (1.9 mg/ml) and P₄₀ (1.6 mg/ml) were recorded and analyzed as described in Materials and Methods. Arrows 1 and 2 denote the spectra of P₄₀ and P₄₁, respectively.

FIG. 6. Thermal unfolding of P₄₀ and P₄₁. The thermal unfolding of P₄₀ and P₄₁ was determined by analysis of CD at 220 nm as described in Materials and Methods. Curves 1 and 2 denote the melting points of P₄₀ and P₄₁, respectively.

Discussion

Analysis of GPR residues essential for catalysis. The removal of ~30% of GPR residues without loss of enzyme activity or the capacity for autoprocessing indicates that the residues essential for GPR catalysis must reside in the remaining 70% of the protein. Comparison of the amino acid sequences of GPR from B. megaterium and B. subtilis has previously failed to reveal any consensus sequence for any known class of protease (32). This analysis can now be extended further (Fig. 7), as the sequence of GPR from Clostridium acetobutylicum has recently become available from Genome Therapeutics Corporation (Waltham, Mass.) through the Internet at www.cric.com. Analysis of these three protein sequences shows that the sequence in the P₄₀→P₄₁ cleavage site is the same in all three proteins, but the clostridial protein has a much shorter propeptide. Again, there are no conserved signature consensus sequences for aspartic or cysteine proteases, metalloproteases, or serine proteases, the four main classes of proteases (2, 19–

Analysis of metal ions in GPR. Previous work has shown that Ca²⁺ is required for P₄₁ stability, although Ca²⁺ is likely not required for catalytic activity (29). However, incubation of P₄₁ as described in Materials and Methods with chelators such as Ophen and HQSA, which have a low affinity for Ca²⁺ but a high affinity for other divalent cations (i.e., Zn²⁺), gave no inhibition of GPR activity (data not shown). These results suggest that GPR is not a metalloprotease, which is consistent with the absence of the classical HEXXH motif of metalloproteases from GPR (21, 32). However, it is still possible that GPR is an unusual metalloprotease with a very tightly bound divalent cation. To test this point directly, we diazylated concentrated GPR solutions against metal-free buffer (except for the stabilizing Ca²⁺ ions) and analyzed the protein for metal ions by atomic absorption spectroscopy. Analysis of both P₄₁ and P₄₀ showed that Ca²⁺, Co²⁺, Cu²⁺, Fe²⁺, Mg²⁺, and Mn²⁺ were present at <0.03 mol/mol of enzyme subunits, while Zn²⁺ was present at <0.05 mol/mol of enzyme subunits (data not shown).
These data effectively rule out the possibility that GPR is a metalloprotease.

**DISCUSSION**

Proteases include enzymes that have been classified into four families—aspartic and cysteine proteases, metalloproteases, and serine proteases—as well as a number of unclassified enzymes (2, 18–21). The data in this communication clearly show that GPR is not a cysteine protease, metalloprotease, or serine protease. Presumably, the inhibition of GPR by PMSF was due to reaction with some amino acid residue other than a serine or with a serine residue essential for protein stability but not catalysis. One catalytic residue other than serine that might react with PMSF is threonine, and recently the catalytic residue in the proteasome has been identified as an N-terminal threonine residue (27). While P41 does not have an N-terminal threonine residue, there are nine threonine residues that are conserved in the three known GPR sequences (Fig. 7). Thus, it is formally possible that GPR could utilize a threonine residue for catalysis. However, the proteasome is inhibited by DCI (4), and this inhibitor had no effect on P41. While the latter finding suggests that GPR is not a threonine protease, it remains possible that the highly specific P41 is a threonine protease; further mutagenesis will be required to test this possibility.

The facts that GPR’s pH optimum is 8 (28) and that the enzyme is not inhibited by pepstatin also suggest that GPR is not an aspartic protease, as these enzymes are generally active at much lower pH values and are inhibited by pepstatin (20). However, it is formally possible that GPR is an aspartic protease that is active at neutral pH, possibly because of the environment around the active-site aspartates, as has been suggested for several aspartic proteases (5, 25), and perhaps the high specificity of GPR’s active site precludes inhibition by pepstatin. Indeed, while GPRs lack the conserved motifs found in many aspartic proteases, the sequences of GPR from the three species examined do contain seven conserved aspartate residues.

**TABLE 2. Enzymatic activities of GPR variants**

<table>
<thead>
<tr>
<th>GPR variant</th>
<th>GPR sp act</th>
</tr>
</thead>
<tbody>
<tr>
<td>P41 S219A</td>
<td>85</td>
</tr>
<tr>
<td>P41 S229A</td>
<td>103</td>
</tr>
<tr>
<td>P41 S274A S276A</td>
<td>70</td>
</tr>
<tr>
<td>P41 C115A</td>
<td>120</td>
</tr>
<tr>
<td>ΔP41</td>
<td>97</td>
</tr>
</tbody>
</table>

*The specific activities of the GPR variants in crude extracts were measured as described in Materials and Methods. Note that all GPR variants were over-expressed to approximately the same level.

**FIG. 7.** Amino acid sequence alignment of GPRs from various species. The protein sequences were aligned by using the ClustalW program. Gaps introduced into the sequences are indicated by dashes. Positions at which identical residues are present in proteins from all three species are indicated by asterisks below the sequence. Numbers at the right indicate the last residue in that row for each protein. The labeled arrows denote the site cleaved in the generation of P41 (a), the site cleaved in generation of P42 (b), the single cysteine in GPR from *Bacillus* species (c), two serine residues conserved in all three proteins (d and e), the likely site of trypsin cleavage generating P30 from P41 (f), and serine residues conserved in the *Bacillus* proteins (g and h). Sources of the sequences: *Bacillus megaterium* (B. meg) (32), *Bacillus subtilis* (B. su) (32), and *Clostridium acetobutylicum* (C. acet) (the Internet at www.cric.com).
served lysine, and five conserved arginines. Although none of these amino acids have been reported to be the direct catalytic residues in proteases, several (e.g., histidine and lysine) participate in catalysis, in particular in serine proteases (19). The lack of inhibition of GPR by NAI suggests that a tyrosine residue does not participate in catalysis (22), but we have as yet no direct proof of the involvement (or lack of involvement) of the other residues. Again, further mutagenesis will be required to examine the role of these other residues and to definitively identify the catalytic residues in GPR.

Recently, a substrate-specific protease has been identified in the obligate intracellular pathogen Chlamydia trachomatis (9). This protease, termed EUO, appears specific for the histone-like proteins which cause chromatin condensation late in this organism’s life cycle, and in this regard EUO displays some similarity to GPR in its biological action. However, EUO differs tremendously from GPR in size, sequence, and inhibition by both pepstatin and a serine protease inhibitor (9).

The present work shows clearly that only the N-terminal two-thirds of P41 and autoprocessing of P46 to P41. This is shown not only by the activity and autoprocessing of the 27- to 30-kDa forms generated from P41 and P46 by trypsin digestion but also by the absence of 30 of the C-terminal residues in the Bacillus GPRs from the clostridial enzyme. However, the C-terminal third of the molecule is needed to impart stability at least to P41. It is striking that the C-terminal regions of both P46 and P41 are removed by trypsin cleavage at the same bond (K268). The trypsin sensitivity of the bond between K268 and E269 suggests that these residues are a part of a structure, possibly a loop, that is exposed to the solvent. The loss of the 102 C-terminal residues from P46 without loss of the autoprocessing activity of P30 further suggests that these C-terminal residues form a domain (or domains) which is distinct from the catalytic domain of the enzyme, although the C-terminal domain may be important in stabilizing the intact enzyme. The resistance of P30 to further trypsin digestion is in contrast to the lability of the 27- and 29-kDa forms generated from P41 by trypsin. The degradation and instability of the latter species, the lack of reactivity of the SH group in P30 (in contrast to the SH group in P41), the much greater trypsin sensitivity of P41 than of P46, the lower melting temperature for P41 than for P46, and the greater susceptibility of P41 than of P46 to unfolding by GuHCl are further evidence for a significant difference in tertiary structure between P46 and P41, with the latter presumably having a less compact structure which is less stable and more accessible to DTNB and trypsin. Such a difference in the structure and stability of zymogen and active enzyme has been seen with a number of other proteases, as the zymogen’s propeptide not only helps maintain the enzyme in an inactive state but also facilitates conversion, a clear challenge for future work will be to elucidate the precise nature of this conformational change.

ACKNOWLEDGMENTS

We are extremely grateful to Douglas Smith of Genome Therapeutics Corporation for providing unpublished genome sequence data. We also thank Donald D. Mucio and Marta Ferrarroni for their help with collecting and interpreting the CD spectra and Barbara Setlow for assistance with some experiments. The AVIV 62DS spectropolarimeter was obtained with funds from NSF grant DIR 8820511. This work was supported by grant GM19698 from the National Institutes of Health (P.S.).

REFERENCES


