Characterization of a Prolyl Endopeptidase from \textit{Flavobacterium meningosepticum}

COMPLETE SEQUENCE AND LOCALIZATION OF THE ACTIVE-SITE SERINE*

(Received for publication, October 17, 1991)

Sylvie Chevallier‡, Pierrette Goeltz‡, Pierre Thibault‡, Denis Banville§, and Jean Gagnon‡

From the ‡Laboratoire de Biologie Structurale, Commissariat à l’Energie Atomique et Centre National de la Recherche Scientifique, URA 1333, DSV/DBMS, CEN-G 85 X, 38041 Grenoble, Cédex, France, the §Institute for Marine Biosciences, National Research Council of Canada, Halifax, B3H 3Z1, Canada and the ‡Biotechnology Research Institute, National Research Council of Canada, Montreal H4P 2R2, Canada

A prolyl endopeptidase was purified from \textit{Flavobacterium meningosepticum}. It was digested with trypsin. Two oligonucleotides, based on tryptic peptide sequences and used in PCR experiments, amplified a 300-base pair (bp) fragment. A 2.4-kilobase EcoRI fragment that hybridized to the 300-bp probe was cloned in \textit{λ}ZAP and sequenced from both strands. It contains a reading frame of 2115 bp, encoding the complete protein sequence of 705 amino acids. Ion-spray mass spectrometry experiments demonstrated the presence of an NH\textsubscript{2}-terminal signal peptide: the periplasmic mature protease is 685 residues in length for a molecular mass of 76,184 Da. The prolyl endopeptidase showed no general sequence homology with known protein sequences except with that of porcine brain prolyl endopeptidase. In order to identify the active-site serine, the prolyl endopeptidase was labeled with [\textsuperscript{3}H]diisopropyl fluorophosphate. One labeled peptide was purified and sequenced. The active-site serine was located in position 536 within the sequence GRSNGGG. This sequence is different from the active-site sequence of the trypsin (GDSGGP) and subtilisin (GTSMAS) families.

Intracellular prolyl endopeptidase appears to be a ubiquitous protease. It is found in various vertebrate tissues such as human uterus (Walter \textit{et al.}, 1971), lamb kidney (Walter, 1976), porcine muscle (Moriyama \textit{et al.}, 1988), and human brain (Kalwant and Porter, 1991). An extracellular prolyl endopeptidase was also purified and characterized in mushrooms: \textit{Lycoperdon cinerascens} (Yoshimoto \textit{et al.}, 1988) and \textit{Agaricus bisporus} (Sattar \textit{et al.}, 1990). Several proteases such as aminopeptidases (prolidase, aminopeptidase P), carboxypeptidases, and dipeptidases show a strict specificity for prolyl residues. But prolyl endopeptidase is the only enzyme that cleaves specifically on the COOH-terminal side of prolyl residues within the sequence X-Pro-Y (X and Y are any amino acids or peptides).

Recently, the sequence of porcine brain prolyl endopeptidase was determined, and the active-site seryl residue was identified by labeling with the inhibitor [\textsuperscript{3}H]diisopropyl fluorophosphate ([\textsuperscript{3}H]DFP) (Rennex \textit{et al.}, 1991), which binds covalently to the activated serine of the catalytic site of enzymes such as serine proteases and serine esterases (Cohen \textit{et al.}, 1967). However, this enzyme showed no homology with type I (trypsin) and type II (subtilisin) serine proteases. The sequence surrounding the active-site seryl residue was different from the known sequences, suggesting that the prolyl endopeptidase belongs to a new type of serine protease.

Among more than 500 microorganisms tested, only \textit{Flavobacterium meningosepticum} showed a post-proline cleavage-like activity (Yoshimoto and Tsuru, 1978). The bacterial prolyl endopeptidase has been purified. Its molecular mass as determined by gel filtration was 74,000 Da, whereas a molecular mass of 76,000 Da was found by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and its pI was estimated at 9.6. The endopeptidase was classified as a serine protease on the basis of its inhibition by DFP (Yoshimoto \textit{et al.}, 1980). It was demonstrated that the enzyme has an absolute requirement for an X-Pro bond in the \textit{trans} configuration immediately preceding the Pro-Y scissile bond (Lin and Brandts, 1983). However, the conformational specificity of vertebrate prolyl endopeptidases has never been investigated.

Because of the unique structure of the prolyl residue, X-Pro bonds have a relatively high intrinsic probability (10-30%) of being in the \textit{cis} rather than in the \textit{trans} configuration (Brandts \textit{et al.}, 1975) as compared with less than 0.1% for peptide bonds implicating other amino acids (Ramachandran and Mitra, 1976). Thus, prolyl residues play an important role in polypeptide chain conformation and in the process of protein folding, as isomerization of the peptidyl-prolyl bond is a rate-limiting step (Brandts \textit{et al.}, 1975).

As prolyl endopeptidase is the only protease that cleaves specifically on the COOH-terminal side of proline, it could become an interesting tool for protein sequencing. Moreover, because of its ability to hydrolyze the Pro-Y bond only if the preceding X-Pro bond is in the \textit{trans} configuration, the prolyl endopeptidase from \textit{F. meningosepticum} could be very useful.

* This work was supported by the Commissariat à l’Energie Atomique, the Centre National de Recherche Scientifique and the National Research Council of Canada. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Biologie Structurale, Commissariat à l’Energie Atomique et Centre National de la Recherche Scientifique, URA 1333, DSV/DBMS, CEN-G 85 X, 38041 Grenoble, Cédex, France. Tel.: 76-88-44-85; Fax: 76-88-51-53.

1 The abbreviations used are: DFP, diisopropyl fluorophosphate; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; Z, benzoyl oxy carbonyl; pNA, p-nitroaniline; 2-NNap, 8-naphthylamide; TPCK, tosylphenylalanyl chloromethyl ketone; TLCK, 1-chloro-3-tosylamido-2-amino-heptane; PCR, polymerase chain reaction; LC-MS, liquid chromatography-mass spectrometry; HPLC, high-performance liquid chromatography; bp, base pair(s); kb, kilobase(s).
for the study of the isomerization and configuration of prolyl-prolyl bonds.

In this paper, we report the DNA sequence encoding the prolyl endopeptidase and the deduced amino acid sequence of the enzyme. About 50% of the predicted amino acid sequence was confirmed by sequencing of numerous tryptic fragments that could be located along the sequence. The deduced molecular mass of the protein was confirmed by ionm mass spectrometry. When compared with the sequence of porcine brain prolyl endopeptidase, the sequence of the bacterial enzyme show more than 38% homology. In order to determine to which type of protease this enzyme belongs, the prolyl endopeptidase from F. meningosepticum was labeled with [3H]DFP and the active site identified as a seryl residue. The bacterial prolyl endopeptidase does not present any general homology with the known families of serine proteases, and to which type of protease this enzyme belongs, the prolyl endopeptidase were collected and their amino acid composition determined.

**Polyacrylamide Gel Electrophoresis—**The SDS-PAGE method (Laemmli, 1970) was systematically used to control the enzyme purity. The gel was stained in Coomasie Brilliant Blue G250 and destained in 10% acetic acid, 30% methanol.

In nonn dentating conditions, the proteins migrated in a 7.5% (w/v) polyacrylamide separating gel (Davis, 1968). The enzyme activity was located in the gel by a modification of the method of Yoshimoto et al. (1980). The gel was incubated for 20 min at 30 °C, in the presence of 1 ml of Z-Gly-Pro-2-NNap (2.25 mg/ml) in 1,4-dioxane and 5 ml of 20 mM phosphate buffer, pH 6.2. A volume of 200 μl of GBC Fast Garnet salt (22.5 mg/ml containing 10% (w/v) Triton X-100 in 1 m acetate buffer, pH 4.0) was mixed with 20 μl of 0.2 M NaNO3. After 10 min at 4 °C, this was added to the mixture above. The incubation was performed for 1 h at 30 °C.

**Amino Acid Analysis and Sequence Determination—**Prolyl endopeptidase that was submitted to amino acid analysis was purified as described above, except that one additional chromatographic step was carried out using a C8 reversed-phase HPLC column (Aquapore RP-300, 7 μm, 100 × 2.1 mm). Samples were hydrolyzed in the vapor of 6 N HCl containing 1% (w/v) phenol for 24, 48, and 72 h at 110 °C. Amino acid analysis was performed using a Beckman Model 6300 analyzer.

Prior to sequence determination the prolyl endopeptidase was reduced and alkylated (Johnson et al., 1980). The enzyme was lyophilized and digested with 2% (w/v) trypsin in 100 mM Tris-HCl, 10 mM CaCl2, pH 8.5, for 18 h at 37 °C. The digest was fractionated on a C8 reversed-phase HPLC column (Aquapore RP-300, 7 μm, 100 × 2.1 mm) using 0.1% (v/v) trifluoroacetic acid and linear gradient from 5 to 65% (v/v) of 70% (v/v) acetonitrile for 45 min at a flow rate of 200 μl/min. Amino acid sequence determination was performed using an Applied Biosystems gas-phase sequencer model 477A with on-line analysis of the phenylthiohydantoin-derivatives. In the case of GPC products, a portion from each sequence cycle was collected and the radioactivity measured for 5 min, after addition of 12.5 ml of liquid scintillation mixture.

**DNA Cloning and Sequencing—**Several degenerate oligonucleotides were synthesized based on the available amino acid sequence of trypptic peptides and were assayed in pair in a polymerase chain reaction (PCR) (Saiki et al., 1985; Mullis and Foalona, 1987) using the AmpliTaq DNA polymerase (Cetus) with bacterial DNA prepared according to Maniatis et al. (1982). One pair of oligonucleotides, PRO-1 (5′ AA(T)/G/C/GG(A/A)A(T)/C(T)/A(C)T/G) and PRO-8 (5′ C/GGG/C/G(T)/A/G/T/A(A)/A/A(A)/G/C) derived, respectively, from the following peptide sequences NPENYV and DFIAAG (re- spectively, 414–419 and 511–516 in Fig. 3), were amplified a fragment of 1 ml of Z-Gly-Pro-2-NNap (2.25 mg/ml) in 1,4-dioxane and 5 ml of 20 mM phosphate buffer, pH 6.2. A volume of 200 μl of GBC Fast Garnet salt (22.5 mg/ml containing 10% (w/v) Triton X-100 in 1 m acetate buffer, pH 4.0) was mixed with 20 μl of 0.2 M NaNO3. After 10 min at 4 °C, this was added to the mixture above. The incubation was performed for 1 h at 30 °C.

**Mass Spectrometry—**All mass spectra were acquired on a SCIEX (Thornhill, Ontario, Canada) atmospheric pressure ionization III
triple-quadrupole mass spectrometer equipped with an atmospheric pressure ionization source and an SCIEX IonSpray™ interface. Instrument control, data acquisition, and data processing were achieved via a Macintosh IIx computer. Cesium iodide clusters (Pleasance et al., 1991) were used to calibrate the mass scale and to adjust the resolution to unity (40% valley definition).

LC-MS analyses used a Hewlett Packard 1090L liquid chromatograph and a 1-mm i.d. × 25-cm Vydac 201TP microbore column coupled directly to the IonSpray™ interface. Separations were achieved using a linear gradient of 5 to 50% acetonitrile (0.1% trifluoroacetic acid) in 29 min and a flow rate of 50 μl/min with no split to the interface. Full-scan analyses (600-1400 Da) were obtained using dwell times of 5 ms per 1-Da step.

In LC-MS-MS experiments the same chromatographic system as that described above was used, and tandem mass spectra were obtained following mass selection of precursor ions by the first quadrupole mass analyzer while scanning the third quadrupole mass analyzer with a dwell time of 5 ms/Da. Collisional activation of energies of 40 eV (laboratory frame of reference) was obtained by introducing argon into the rf-only quadrupole at a pressure sufficient to attenuate the precursor ion beam by 80-90%. This attenuation corresponded to a target gas thickness of typically 4 × 10⁻¹⁴ molecules/cm².

Lonsprays mass spectra of proteins were obtained by injecting 5 μl of a dialyzed protein solution (0.6 mg/ml in 10% acetic acid), into a stream of 50% acetonitrile (0.1% trifluoroacetic acid) delivered by the HPLC pump at a flow rate of 5 μl/min. Coherent series of multiply-protonated ions were used to determine the molecular weights of the proteins and to reconstruct their molecular weight profiles (Mazn et al., 1989). The mass spectrometer was calibrated over a narrow mass range of 700 Da with a step size of 0.5 Da and a dwell time of 5 ms per step. The mass accuracy was first checked by running a standard solution of 1 mg/ml of bovine carbonic anhydrase (Sigma Biochemicals), and mass assignment was typically within 0.01% of the expected isotope-averaged molecular weight calculated from the known sequence.

Active-site Labeling with [H]DFP—A 30-fold molar excess of [H]DFP (4 Ci/mm) was added to 1 nmol of prolyl endopeptidase. After 15 min at room temperature, labeling was stopped with a large excess of unlabeled DFP (7.5 μmol) in 2-propanol. The labeled enzyme was dialyzed at 4 °C against 10 mm ammonium bicarbonate. The solution was concentrated (in a Speed-Vac concentrator) and the endopeptidase was treated with cyanogen bromide (Gross, 1967) at an endopeptidase/CNBr ratio of 1/150 (w/w). The reaction was performed in 70% formic acid for 24 h at 4 °C under nitrogen and in the dark. The digest was fractionated on a C₄ reversed-phase HPLC column (Aquapore RP-300, 7 μm, 100 × 2.1 mm) using 0.1% (v/v) trifluoroacetic acid in acetonitrile for 45 min at a flow rate of 200 μl/min. Peaks, detected at 215 nm, were collected manually. For each fraction, the radioactivity of an aliquot (20 μl) was measured for 5 min, after addition of 7.5 ml of liquid scintillation mixture. The fraction containing the majority of the radioactivity was repurified on a C₄ reversed-phase HPLC column (SPHERI-5 RP-18, 5 μm, 100 × 2.1 mm) using the same gradient, and the radioactivity of each fraction was measured.

Expression of Prolyl Endopeptidase in Escherichia coli—A 3-ml preculture was started from a single colony of DH5α cells transformed with the plasmid pBS-EcoRI-#4 (containing the 2.4-kb EcoRI fragment) and grown in L broth containing 100 μg/ml ampicillin (overnight at 37 °C). Two cultures (500 ml of L broth containing 100 μg/ml ampicillin) were started from this preculture and were grown for 22 h at 37 °C. Samples of 50 ml were removed at various times. The cells were harvested by centrifugation (8000 rpm, for 30 min, at 4 °C) and the medium was assayed for enzyme activity toward Z-Gly-Pro-pNA. The periplasmic proteins were removed by a conventional osmotic shock procedure and by washing the resulting pellet with ice-cold 1 M Tris·HCl, pH 7.4. The intracellular proteins were removed by sonication of the pellet. For each time point, both fractions were assayed for enzyme activity toward Z-Gly-Pro-pNA. As control, the same conditions and assays were used with nontransformed bacteria.

RESULTS

Purification of the Endopeptidase—Prolyl endopeptidase was purified 250-fold from cultures of F. meningosepticum. Starting from 60-g wet weight of cells, 0.8 mg of prolyl endopeptidase were obtained with a specific activity of 65 units/mg and a yield of 12%.

Purity of the Enzyme—After Mono S chromatography, the enzyme preparation showed one major band using SDS-PAGE. This band migrated with an apparent molecular mass of 72,000 Da. Using polyacrylamide gel electrophoresis in nondenaturing conditions, one major band was observed that showed activity toward Z-Gly-Pro-2-NNap, as described under “Experimental Procedures.”

Endopeptidase Activity toward Synthetic Peptide—Prolyl endopeptidase activity was assayed using the synthetic peptide EYYPDNYLRT. The peptides obtained from the digest were purified on a C₄ reversed-phase HPLC column (Fig. 1). Their amino acid compositions were determined. The results showed that the Pro-Asn bond was the only one cleaved. This confirms that the purified endopeptidase shows specificity toward prolyl residues.

Endopeptidase Sequencing—The native protein was resistant to Edman degradation, suggesting that the NH₂-terminal residue was blocked. Treatment of the enzyme with 1.5 n HCl (Clarke and Witkowski, 1981) did not enable protein sequencing (data not shown), indicating that the endopeptidase was not blocked by NH₂-terminal formylmethionine. To obtain information about the amino acid sequence, the endopeptidase was reduced, carboxymethylated, and digested with trypsin. The digest was chromatographed on a C₄ reversed-phase HPLC column and several of the purified peptides were submitted to sequence analysis. The sequences as well as the location of these peptides within the polypeptide chain are shown in Fig. 3.

DNA Cloning—Several oligonucleotides were synthesized, based on the amino acid sequences of tryptic peptides. These primers were made in both orientations and used in pairs for PCR on bacterial DNA. One combination of two oligonucleotides (see “Experimental Procedures”) amplified a 300-bp fragment. The sequences of some tryptic peptides could be located within the deduced amino acid sequence, confirming that this fragment encoded part of the endopeptidase.

The bacterial DNA was cleaved with different restriction enzymes (BamHI, HindIII, and EcoRI) and analyzed by hybridization with the amplified 300-bp fragment. A 2.4-kb EcoRI fragment was found to hybridize to the probe and was cloned from a size-selected DNA library of F. meningosepticum DNA in λZAP. The 2.4-kb insert was sequenced in both orientations (Fig. 2). It contains a single open reading frame beginning at the first 5′-initiator codon and encoding a protein of 705 amino acids, for a calculated molecular mass of 78,840 Da (Fig. 3). More than half of the deduced amino acid sequence of the enzyme was confirmed by the analysis of numerous tryptic peptides located along the predicted sequence.

Analysis of the amino acid sequence reveals the presence of a hydrophobic NH₂-terminal region (20 to 1 in Fig. 3), indicating the presence of a potential signal peptide. As a result, the mature endopeptidase is predicted to contain 685 amino acid residues for a molecular mass of 76,784 Da. The amino acid composition calculated from the predicted sequence of the mature enzyme correlates well with the composition determined after acid hydrolysis of the isolated endopeptidase (data not shown). Prolyl endopeptidase contains about 10% of lysyl residues. This proportion is very high in prokaryotic proteins (Doolittle, 1986). The pl calculated from the amino acid composition (deduced from the DNA sequence) is 9.6, similar to the value determined experimentally by Yoshimoto et al., in 1980.

Mass Spectrometry—Confirmation of the prolyl endopeptidase.
Characterization of a Bacterial Prolyl Endopeptidase

that the molecular mass of this complex is 76,944 ± 8 Da, consistent with the covalent binding of single DFP group (164 Da) on the protein.

Confirmation of the NH2-terminal amino acid sequence was obtained by LC-MS analysis of a tryptic digest of the prolyl endopeptidase. In this latter experiment singly- and doubly-protonated ions at m/z 773,387, and 791,396 Da were observed for the pyroglutamate and free glutamine forms of tryptic peptide 1–7, respectively. As expected, these two peptides showed similar retention characteristics on the reversed-phase column although the intensity of the former ions (pyroglutamate at position 1) was larger than the latter ones (glutamine in position 1). The fact that these two forms were observed in the tryptic digest suggests that deamidation of the glutamine residue might have occurred during storage of the digest under acidic conditions (0.1% trifluoroacetic acid in water). Considering that the mass accuracy with ionspray is typically ±0.01%, the possibility of having glutamine at position 1 is further substantiated by the good agreement of the measured molecular mass (76,782 ± 8 Da) compared with that of the predicted mass of the prolyl endopeptidase (76,783 Da, if the 2 cysteines form a disulfide bridge). If pyroglutamate had originally been in position 1, the corresponding molecular mass would be 18 Da lower, well outside the experimental uncertainty.

Active-site Serine Labeling with [3H]DFP—Prolyl endopeptidase was labeled with [3H]DFP and cleaved with cyanogen bromide. The resulting peptides were purified on a C18 reversed-phase HPLC column, and the radioactivity of each fraction was measured (Fig. 5). After chromatography, the recovery of label was about 91%. The radioactivity was distributed over two peaks, but the fraction 19 contained most of the label. The incorporation of [3H]DFP into the labeled peptide (contained in this fraction) was estimated at 1.14 mol of DFP/mol of peptide. Thus the fraction 19 was repurified on a C18 reversed-phase HPLC column. The chromatography revealed a main radioactive peak. This major radioactively labeled peptide was sequenced. The single amino acid sequence ALSGRXNGG (X represents an unidentified residue) was obtained with yields in picomoles shown in parentheses: A (41), L (23), S (7), G (10), R (2), X, N (3), G (3), G (3).

This peptide was located in the sequence of the prolyl
FIG. 3. Complete nucleotide sequence of the 2.4-kb EcoRI fragment and corresponding amino acid sequence of prolyl endopeptidase. The nucleotide sequence (top line) is numbered 1 to 2408 from the first nucleotide of the EcoRI site. The predicted amino acid sequence (bottom line) is shown in the single-letter code. Numbers —20 to —1 of the protein sequence indicate the potential signal peptide. Number +1 of the amino acid sequence indicated by Δ, is assigned to the NH₂-terminal residue of the mature enzyme. The sequence of tryptic peptides is underlined.

![Graph A](image1.png)

**FIG. 4.** Ionspray mass spectra of the mature prolyl endopeptidase (A) and the prolyl endopeptidase inhibited by DFP (B). Insets represent the reconstructed molecular profile calculated from the multiply-charged ions series. Both spectra were obtained with the consumption of approximately 3 μg of protein.

endopeptidase and corresponded to amino acids 531–539.

The radioactivity was measured at each sequencing cycle. It was detected at cycle 6, but the amino acid in this position could not be identified. However, based on the known sequence, this result indicates that this residue corresponded to Ser-536 and was labeled with [³H]DFP. Moreover, the region surrounding this seryl residue is very similar to the region (positions 549–557) of the porcine brain prolyl endopeptidase (Rennex et al., 1991) which has been shown to contain the active site. These results show that the active-site serine of the prolyl endopeptidase from *F. meningosepticum* is located in position 536.

**Expression of Prolyl Endopeptidase—** *E. coli* DH₅α were transformed with pBS-EcoRI-#4 and grown for 1991 h at 37 °C in L broth in the presence of ampicillin. In the case of periplasmic proteins from the supernatant obtained after osmotic shock, the specific activity reached a plateau after 14 h, and stayed constant to 20 h of culture. The activity toward the peptide Z-Gly-Pro-pNA correlates well with the intensity of a new protein migrating at the level of the purified prolyl endopeptidase in SDS-PAGE (data not shown).

After 20 h of culture and precipitation by ammonium sulfate (65–90% saturation) of the total cellular extract obtained by sonication, the prolyl endopeptidase activity was assayed toward the synthetic peptide EYYDPNYLRT. The peptides surrounding this seryl residue is very similar to the region (positions 549–557) of the porcine brain prolyl endopeptidase (Rennex et al., 1991) which has been shown to contain the active site. These results show that the active-site serine of the prolyl endopeptidase from *F. meningosepticum* is located in position 536.
plasmic localization. Band, migrating at the level of the native protein, showed the prolyl endopeptidase from F. meningosepticum with a yield of 12% using a procedure similar to that of Yoshimoto et al. (1980). The [3H]DFP-endopeptidase was treated with cyanogen bromide, and the digest was chromatographed on a C4 reversed-phase HPLC column using the gradient described under "Experimental Procedures." The radioactivity of an aliquot (20 μl of each fraction was measured (shaded boxes). Fraction 19 contained most of the label.

DISCUSSION

Prolyl endopeptidase from F. meningosepticum was purified with a yield of 12% using a procedure similar to that of Yoshimoto et al. (1980). Its apparent molecular mass was estimated at 72,000 Da by SDS-PAGE. Using polyacrylamide gel electrophoresis in nondenaturing conditions, one major band, migrating at the level of the native protein, showed activity toward Z-Gly-Pro-2-NNap, as described under "Experimental Procedures."

The specificity of the prolyl endopeptidase was assayed toward the synthetic peptide EYYDPYNYLRT. The peptides generated were purified by reverse-phase HPLC (Fig. 1) and their amino acid composition was determined. Amino acid analysis revealed that only the Pro5-Asn5 bond was cleaved. This result indicated that, in the case of this peptide, the endopeptidase was specific for prolyl residues, like the prolyl endopeptidase purified and characterized by Yoshimoto et al. (1980). This method could be used to investigate specificity using other synthetic peptides.

The following inhibitors were tested for their effect on the enzyme activity toward the peptide Z-Gly-Pro-pNA (data not shown): EDTA, soybean trypsin inhibitor, TLCK, TPCK, pepstatin, phenylmethanesulfonyl fluoride, and DFP. Among these, only DFP showed complete inhibition of the prolyl endopeptidase activity, as previously shown by Yoshimoto et al. (1980). This compound is known to specifically inhibit serine proteases by covalent binding to the active-site seryl residue. Although the prolyl endopeptidase was not sensitive to phenylmethanesulfonyl fluoride, another serine protease inhibitor, inhibition by DFP indicated that the enzyme had an active-site serine, essential for its catalytic activity.

With a view to studying the primary structure of the prolyl endopeptidase, the reduced and alkylated protein was digested with trypsin and several of the purified tryptic peptides were submitted to sequence analysis. Several degenerate oligonucleotides were synthesized based on available peptide sequences, and used as primers in PCR experiments. One pair of these primers, PRO-1 and PRO-8, respectively, based on the sequences 511DFIAAG516, amplified a 300-bp DNA fragment that was later used as a probe. A 2.4-kb fragment, resulting from an EcoRI digest of F. meningosepticum DNA, hybridized to this probe. This EcoRI fragment was purified, cloned in λZAP, and sequenced in both orientations (Fig. 2). The fragment revealed a single reading frame of 2115 bp beginning at the first 5' initiator codon (ATG) in position 50 and ending at a stop codon (TAG) in position 2165 as indicated in Fig. 2.

The sequence of the prolyl endopeptidase, deduced from the DNA sequence, consists of 705 amino acid residues (20 to 685 in Fig. 3). More than 50% of the predicted sequence is confirmed by the sequences of the tryptic fragments as indicated in Fig. 3. Analysis of the amino acid sequence reveals the presence of a hydrophobic region at the NH2-terminal end. The computer-predicted secondary structure of this segment is that of an α-helix followed by a β-turn (Wu, 1986). It can be divided into three regions: a hydrophilic and positively charged NH2-terminal region containing 3 basic residues, a central hydrophobic core, and a polar COOH-terminal end enriched in amino acids with small side chains immediately preceding the presumed cleavage site. Moreover it contains a COOH-terminal alanine residue found in most prokaryotic signal sequences (Gascuel and Danchin, 1986). These results suggested the presence of a signal peptide of 20 amino acid residues (20 to -1 in Fig. 3) preceding the mature enzyme. Thus the mature endopeptidase, 685 amino acids in length, is thought to contain a NH2-terminal glutamyl residue (position 1 in Fig. 3).

As the enzyme appeared by sequence analysis to have a blocked NH2 terminus, the presence of a NH2-terminal pyroglutamyl residue was expected. The value of the molecular mass calculated from the deduced amino acid sequence is

Characterization of a Bacterial Prolyl Endopeptidase

8197
two molecules are represented by a prolyl endopeptidase was determined to be terminal end.

the pyroglutamate form.

bacterial enzyme was compared with the sequence of porcine brain and Wunsch (1970). The predicted sequence of the mature cates that the mature prolyl endopeptidase from man and Wunsch, 78,840

dopeptidase activity toward the peptide Z-Gly-Pro-pNA was

plasmic location of the mammalian prolyl endopeptidases. In

detected in the periplasmic fraction, indicating that the prolyl endopeptidase was indeed located in the periplasm. 1991).

As a consequence, the presence of this signal peptide indi-

Using ionspray mass spectrometry, the molecular mass of

This result confirms that a signal peptide is cleaved to generate the mature enzyme starting with a free glutamine residue at the NH2-terminal end.

As a consequence, the presence of this signal peptide indicates that the mature prolyl endopeptidase from F. meningosepticum is probably periplasmic, in contrast with the cytoplasmic location of the mammalian prolyl endopeptidases. In order to confirm this periplasmic location in F. meningosepticum, experiments were carried out to extract the periplasmic proteins with a conventional osmotic shock procedure. Endopeptidase activity toward the peptide Z-Gly-Pro-pNA was detected in the periplasmic fraction, indicating that the prolyl endopeptidase was indeed located in the periplasm.

The predicted amino acid sequence of the prolyl endopeptidase from F. meningosepticum does not reveal any homology with other known proteins, except with that recently determined for porcine brain prolyl endopeptidase (Rennex et al., 1991). In Fig. 6, the amino acid sequence of the prolyl endopeptidase from F. meningosepticum is aligned (Needleman and Wunsch, 1970) and compared with the sequence of the porcine brain enzyme. The two sequences share 38% identical residues, the NH2-terminal and COOH-terminal regions being the most homologous.

The prolyl endopeptidase from F. meningosepticum was inhibited by DFP. Mass spectral analysis determined that the molecular mass of the DFP-inhibited enzyme was 76,944 ± 8 Da (Fig. 4B). This correlates with the covalent binding of a single molecule of DFP per molecule of protein. Although the enzyme was inhibited by DFP, indicating that it contained an active-site seryl residue, no general homology was found with serine proteases, except with the porcine brain prolyl endopeptidase. In order to determine the location within the sequence of the active-site seryl residue, the prolyl endopeptidase was labeled with [3H]DFP, cleaved by cyanogen bromide, and the major radioactively labeled peptide was purified (Fig. 5). The radioactivity incorporated in the peptide corresponded to 1.14 mol of DFP per mol of peptide, confirming the result obtained by ionspray mass spectrometry. The labeled peptide was sequenced and located at position 531–539 (Fig. 3). The localization of the incorporated radioactivity within this sequence demonstrates that the active-site serine is located at position 536 and corresponds to the active-site serine at position 554 identified by Rennex et al. (1991) in the case of porcine brain prolyl endopeptidase (Fig. 6).

Concerning the other 2 amino acid residues forming the catalytic triad of serine proteases, only the histidine has clearly been identified in the porcine brain prolyl endopeptidase (His-680) (Stone et al., 1991). The region surrounding this amino acid shows homology to the equivalent sequence in the prolyl endopeptidase from F. meningosepticum, indicating that the histidyl residue in position 675 of the bacterial enzyme may be implicated in the active site (Fig. 6). This remains to be confirmed experimentally. The position of the third amino acid implicated in the catalytic triad, an aspartyl or a glutamy1 residue, is still unknown.

In preliminary experiments on the expression of the prolyl endopeptidase in E. coli, a new enzyme activity toward Z-Gly-Pro-pNA was detected in E. coli transformed by the plasmid pBS-EcoRI-#4 (containing the 2.4-kb EcoRI fragment). It correlates with the production of a novel periplasmic protein that has the same apparent molecular mass as the prolyl endopeptidase from F. meningosepticum in SDS-PAGE (data not shown). The assay toward a synthetic peptide demonstrates that this protein encoded by the plasmid pBS-EcoRI-#4 has the same specificity for prolyl residues as the prolyl endopeptidase naturally produced by F. meningosepticum. These results suggest that the enzyme produced by transformed E. coli is identical to the prolyl endopeptidase from F. meningosepticum.

As the prolyl endopeptidase from F. meningosepticum cleaves specifically on the COOH-terminal side of proline, it could become an interesting tool for protein sequencing, made easily available through expression in E. coli. Furthermore, because of its specificity toward X-Pro bonds in the trans conformation, the bacterial enzyme could also be useful for the study of isomerization and conformation of peptide-tyl-prolyl bonds (Brandts and Lin, 1986).

Acknowledgments—We are indebted to G. Hudry-Clergeon for helpful discussions. The excellent technical help of N. Scherrer with protein sequencing is gratefully acknowledged. We wish to thank G. Arlaud and I. Bally at the Département de Biologie Moléculaire Structurale, Grenoble, for providing us with the synthetic peptide.

REFERENCES

Fig. 6. Amino acid alignment between prolyl endopeptidase from F. meningosepticum and that of porcine brain. The alignment was performed using AALIGN by the method of Needleman and Wunsch (1970). The predicted sequence of the mature bacterial enzyme was compared with the sequence of porcine brain enzyme (Rennex et al., 1991). The conserved residues between the two molecules are represented by a dot.

78,840 Da. Because of the loss of a signal peptide (2056 Da), the mature prolyl endopeptidase would have a molecular mass of 76,784 Da for the free glutamine form or of 76,766 Da for the pyroglutamate form.

Using ionspray mass spectrometry, the molecular mass of prolyl endopeptidase was determined to be 76,782 ± 8 Da (Fig. 4A), about 2056 Da below the deduced mass from the nucleotide sequence (−20 to 685 in Fig. 3). This result confirms that a signal peptide is cleaved to generate the mature enzyme starting with a free glutamine residue at the NH2-terminal end.

In preliminary experiments on the expression of the prolyl endopeptidase in E. coli, a new enzyme activity toward Z-Gly-Pro-pNA was detected in E. coli transformed by the plasmid pBS-EcoRI-#4 (containing the 2.4-kb EcoRI fragment). It correlates with the production of a novel periplasmic protein that has the same apparent molecular mass as the prolyl endopeptidase from F. meningosepticum in SDS-PAGE (data not shown). The assay toward a synthetic peptide demonstrates that this protein encoded by the plasmid pBS-EcoRI-#4 has the same specificity for prolyl residues as the prolyl endopeptidase naturally produced by F. meningosepticum. These results suggest that the enzyme produced by transformed E. coli is identical to the prolyl endopeptidase from F. meningosepticum.
Characterization of a Bacterial Prolyl Endopeptidase


Characterization of a prolyl endopeptidase from Flavobacterium meningosepticum. Complete sequence and localization of the active-site serine. 
S Chevallier, P Goeltz, P Thibault, D Banville and J Gagnon


Access the most updated version of this article at http://www.jbc.org/content/267/12/8192

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/267/12/8192.full.html#ref-list-1