Amino Acid Sequences of Metalloendopeptidases Specific for Acyl-Lysine Bonds from *Grifola frondosa* and *Pleurotus ostreatus* Fruiting Bodies*

(Received for publication, December 19, 1996, and in revised form, August 8, 1997)

Takashi Nonaka‡§, Naoshi Dohmae*, Yohiichi Hashimoto‡, and Koji Takio¶

From the ©Department of Biochemistry, Faculty of Science, Saitama University, Urawa, Saitama 338 and ¥Division of Biomolecular Characterization, The Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-01, Japan

The complete amino acid sequences of two lysine-specific zinc metalloendopeptidases (EC 3.4.24), *Grifola frondosa* metalloendopeptidase (GFMEP) and *Pleurotus ostreatus* metalloendopeptidase (POMEP), from the fruiting bodies of these two edible mushrooms have been established based on the sequence information of the peptides generated from the reduced and alkylated GFMEP and POMEP by proteolytic digestions using GFMEP, trypsin, and other proteases as well as by several chemical cleavages. From the sequences, it was found that GFMEP and POMEP were polypeptides composed of 167 and 168 amino acid residues, from which their molecular weights were calculated to be 18,040.5 and 17,921.3 in accord with the observed (M+H)+ values of 18,028 and 17,927, respectively, as determined by matrix-assisted laser desorption ionization-time of flight mass spectrometry. Two disulfide bonds in GFMEP were found to link Cys5 to Cys75 and Cys77 to Cys97. A nunn- trix-assisted laser desorption ionization-time of flight mass spectrometry. Two disulfide bonds in GFMEP were found to link Cys5 to Cys75 and Cys77 to Cys97. An unusual post-translational modification of GFMEP was corroborated to be a partial attachment of a single man- nose to Thr42. Comparison of the sequences revealed that overall identity between the enzymes was 61.3%. Although a highly homologous sequence was not found in sequence data bases except for a consensus zinc-bind- ing sequence, HEXXH, both metalloendopeptidases somewhat resembled a family of metalloproteases catego- rized as deuterolysin. These proteases together with GFMEP and POMEP do not have conserved third and/or fourth liganding amino acid residues seen in metzincin or thermolysin superfamily proteins and belong to a novel zinc metalloendopeptidase superfamily.

Only a few metalloendopeptidases have been purified and characterized from the culture fluids or fruiting bodies of such higher basidiomycetes as *Flammulina velutipes*, *Lentinus edodes*, and *Ganoderma lucidum* (1, 2). They are sensitive to EDTA and a specific inhibitor, talopeptin, and may be grouped into neutral proteinases owing to their pH optima over a range of 6.7–8.5 and specificity toward the peptide bonds involving the amino group of lysine residues. The unique specificity of the enzyme was confirmed by its application to sequence analyses of some proteins (4, 5). Recently, purifying a MEP (GFMEP) from *Grifola frondosa* fruiting bod- ies, we have characterized this enzyme with a molecular mass of 20 kDa to contain 1 atom of zinc/molecule. This enzyme has several prominent properties including a high thermostability, a maximal activity in an alkaline range of pH 9–10, binding capacity to β-glycans, and specific cleavage of acyl-lysine bonds (6). From the fruiting bodies of another edible mushroom, *Pleu- rotus ostreatus*, a zinc MEP named POMEP has been purified that closely resembles GFMEP in its properties such as molec- ular mass (19 kDa), tolerance to heat and denaturing reagents, and specificity to acyl-lysine bonds, except for its slightly acidic optimum pH (7). These findings led us to postulate that the lysine-specific MEPS might be inherently evolved in basidiomy- cetes, because no protease with such a strict specificity to- ward acyl-lysine bonds has been reported in eumycetes. In this paper, we report the determination of the complete amino acid sequences of the zinc MEPS as the first step to analyze the three-dimensional structure, which may provide the keys for elucidating the structural features responsible for the zinc coordinating catalytic sites as well as for the unique specificity toward acyl-lysine bonds. The data obtained in this work pres- ent valuable sequence information for the structural relation of these MEPS to various zinc metalloproteases, which have recently been grouped into some families on the basis of their amino acid sequences and structural features around the zinc-binding sites (8).

**EXPERIMENTAL PROCEDURES**

**Materials**—Fruiting bodies of the popular edible mushrooms, *G. frondosa* and *P. ostreatus* were purchased from producers in Niigata and Gunma Prefecture, respectively, and stored at −40 °C. The sources of other materials and chemicals used in this work were as follows: TPCK- treated trypsin and TLCK-treated chymotrypsin from Cooper Bio- medical (Malvern, PA); *Staphylococcus aureus* V8 protease from ICN Bio- medicals (Costa Mesa, CA); endoproteinas Aep-N from Boehringer Mannheim (Mannheim, Germany); jack bean α-mannosidase and ther-

© 1997 by The American Society for Biochemistry and Molecular Biology, Inc.

**THE JOURNAL OF BIOLOGICAL CHEMISTRY** Vol. 272, No. 48, Issue of November 28, pp. 30032–30039, 1997

Printed in U.S.A.
molysin from Seikagaku Kogyo (Tokyo, Japan); cyanogen bromide, 2-aminopyridine, and tri-\(n\)-butyl phosphine from Wako Pure Chemical (Osaka, Japan); 4-vinylpyridine and 2-nitro-5-thiocyanatobenzoic acid from Tokyo Kasei (Tokyo, Japan); 2-bromoethyamine hydrobromide from Osaka, Japan); 4-vinylpyridine and 2-nitro-5-thiocyanatobenzoic acid from Kanto Chemical (Tokyo, Japan); BNPS-skatole from Pierce Chemical (Rockford, IL); borodicimethylamine complex from Aldrich Chemical (Milwaukee, WI). Other chemicals were of analytical grade.

**Preparation of S-Alkylated MEP—GFMEP and POMEP were purified as reported from G. frondosa (6) and P. ostreatus (7) fruiting bodies, respectively. Reduced, and pyridylethylated (PE) as described previously (6).**

**Chemical Cleavages—** Methionyl bonds were cleaved according to the method of Gross (10) with 1% (w/v) CNBr in 70% (v/v) formic acid. Tryptophanyl bonds were cleaved with 2-nitro-5-thiocyanatobenzoic acid by the method of Stark (12). Peptide bonds involving asparaginyl residues were hydrolyzed in 2% (v/v) formic acid at 108 °C for 45 min according to Inglis (13).

**Enzymatic Cleavages—** Proteolytic digestion by GFMEP, TPCK-treated trypsin, TLCK-treated chymotrypsin, endoproteinase Asn-N, *S. aureus* V8 protease or thermolysin was carried out at 37 °C in 0.1 M ammonium bicarbonate, pH 9.5, or in 50 mM Tris-HCl, pH 7.5, or 8.0 buffer containing 2 or 4 mM urea (total volume, 80 µl) for 18 h at an enzyme to substrate ratio of 1:50 to 1:20 (mol/mol) unless stated otherwise. PE-GFMEP (2 nmol) was dissolved in 50 mM sodium phosphate buffer, pH 8.0, containing 2 mM urea (total volume, 80 µl) and incubated with endoproteinase Asp-N ([E/S = 1:50, mol/mol] at 37 °C for 18 h. AE-POMEp (100 pmol) was subjected to digestion with GFMEP ([E/S = 1:20, mol/mol] in 0.1 M ammonium bicarbonate containing 2 mM urea (total volume, 80 µl) at 37 °C overnight followed by incubation at 37 °C overnight with TPCK-treated trypsin ([E/S = 1:50, mol/mol] added to the reaction mixture. To determine the disulfide bond pairs, cyanogen bromide-cleaved native GFMEP (3 nmol) was dissolved in 8 µl urea containing 1 mM iodoacetic acid (20 µl) and incubated at 37 °C in the dark for 1 h. The mixture was digested successively with TPCK-treated trypsin and TLCK-treated chymotrypsin in an enzyme to substrate ratio of 1:20 (mol/mol) each in 0.1 M ammonium bicarbonate containing 2 mM urea (total volume, 80 µl) at 37 °C for 18 h. A mannose-containing peptide, T2 (100 pmol), generated from PE-GFMEP by tryptic cleavage was digested with jack bean α-mannosidase (27 units) in 50 mM sodium acetate buffer, pH 4.5, (100 µl) at 30 °C for 20 h, followed by another 20-h incubation with an addition of fresh enzyme.

**Separation of Peptides—** Peptides were first separated by GPC-HPLC on tandem columns of TSK G2000SWXL and TSK G3000SWXL (7.8 x 300 mm each, Tosoh, Tokyo) or a single column of TSK G2000SWXL using a Gilson model 302 pump and a Hewlett Packard HP 1040M diode-array detection system. Elution was conducted with 10 mM phosphate buffer, pH 6.0, containing 6 M guanidine HCl at a flow rate of 0.4 ml/min, and the effluent was monitored at 215, 260, 275, and 290 nm. Further separation of these peptides or all enzymatic digest mixtures were achieved by RP-HPLC using a Gilson HPLC system or a Hewlett Packard model 1090M liquid chromatograph on an Aquapore RP-300 (2.1 x 30 or 4.6 x 100 mm, Applied Biosystems, CA), a Supersphere Select-B (2 x 119 mm, E. Merck, Darmstadt), or a Pico Tag (4.6 x 150 mm, Millipore Corp., MA) column. Elution of peptides was performed with a linear gradient of 0.09% (v/v) trifluoroacetic acid to 64% (v/v) acetonitrile in 0.078% (v/v) trifluoroacetic acid at a flow rate of 0.5 or 0.2 ml/min. The effluent was monitored at 215, 254, 270, and 295 nm with a diode array detector.

**Amino Acid, Sequence, and Sugar Analyses—** Compositional analyses were performed by precolumn derivatization with a Waters Pico Tag system (14). The samples (50–200 pmol) were dried and hydrolyzed in vapor phase of 6 N HCl containing 0.1% (w/v) phenol at 110 °C for 20 h. Automated Edman degradation was carried out with an Applied Biosystems model 477A protein sequencer connected on line to a model 120A PTH analyzer (15) using an in-house generated gas phase sequencing program. Sugar in GFMEP was identified by the method of Suzuki et al. (16).

**Mass Spectrometry—** Mass spectrometric analyses of peptides were performed by capillary liquid chromatography/Frit-FAB MS on a JMS-LX1000 mass spectrometer (JEOL, Tokyo, Japan) equipped with an on-line micro gradient system (Applied Biosystems). Peptides were separated on a capillary column, Fuc18 (0.32 x 250 mm, LC Packings, Zürich) in 0.09% (v/v) trifluoroacetic acid containing 1% (w/v) glycerol with a linear gradient of 0–64% (v/v) acetonitrile for 16 min at a flow rate of 3.5 µl/min. Peptides were monitored with a UV detector model 441 (Millipore) equipped with a capillary flow cell (LC Packings) and directly introduced into a JMS-LX1000 with a Frit probe, which was operated at 2 or 3 kV of accelerating voltage in a mass range of 10–2,200 or 10–1,500 m/z, respectively, and the temperature of the ionization chamber was maintained at 50 °C. MALDI-TOF MS analyses of the whole protein and selected peptides were performed on REFLEX (Bruker-Franzen Analytik, Bremen, Germany) or a Compact MALDI III (Kratz Analytical, Manchester, UK) with ʻcyan-4-hydroxycinnamic acid or sinapinic acid as a matrix.

**Nomenclature of Peptides—** Peptides are designated by a serial number prefixed with a letter. The letters indicate the type of digestion: K, GFMEP; T, trypsin; C, chymotrypsin; D, endoproteinase Asn-N; E, S. aureus* V8 protease; TH, thermolysin; M, cyanogen bromide; W, BNPS-skatole; CT, chymotrypsin and trypsin; TD, trypsin and dilute acid treatment; TCN, tryptic digestion of S-cyanlated protein; TKAE, tryptic and GFMEP digestion of S-aminoethyalted protein. The numbers in the peptide designation do not correspond to the order of their elution in HPLC but rather to their positions in the protein sequence, starting from the amino terminus.

**RESULTS**

**Sequence Analysis of GFMEP—** The overall strategy used for the determination of the complete amino acid sequence of GFMEP is summarized in Fig. 1. More than 80% of the sequence has been determined by Edman degradation, and the results are shown in one-letter codes for the sequence determination. (top lines). The observed molecular mass values by Frit-FAB or MALDI-TOF MS are indicated in parentheses. Within each specified sequence, upper-case letters are proven by Edman degradation (regular type) or Frit-FAB MS (italic type). Lowercase letters denote tentative identification of phenylthiohydantoin. Sequences not identified are indicated by dashes or arrows for a long unidentified sequence. Disulfide bonds link Cys5 to Cys75 and Cys77 to Cys97, respectively. Partial attachment of mannose is indicated by a lower-case M.
Amino Acid Sequences of Metalloendopeptidases

sequence was established by automated Edman degradation of the fragments arising from PE-GFMEP by digestion with GFMEP or trypsin. The sequences of the remainders and the overlaps covering the major fragments were determined based on the sequence data of the peptides derived from two CNBr fragments, M1 and M2+3, by digestion with S. aureus V8 protease as well as those of W3 and D4 generated from PE-GFMEP by BNPS-skatole and endoproteinase Asp-N. The molecular weights of GFMEP, PE-GFMEP, and selected peptides were determined by MALDI-TOF MS and/or Frit-FAB MS to confirm the sequences obtained by automated Edman degradation. The amino-terminal sequence analysis of intact PE-GFMEP established the first 26 residues except for some ambiguities for four consecutive Ala residues.

Cleavage with GFMEP—As seen in Fig. 2A, six peptides, K1, K1', K2, K3, K4+5, and K4+5' were resolved from a GFMEP digest of PE-GFMEP (2 nmol) by RP-HPLC on an Aquapore RP-300 column. Peptides K2 and K4+5 were isolated by rechromatography. Lacking a lysine residue in the composition (Table I), large peptides K1 and K1' must be the amino-terminal fragments, in which all four Cys residues and all Trp residues in GFMEP are contained as indicated by the absorptions at 254 (PE-Cys) and 290 nm (Trp) in the elution profile. Amino acid compositions of peptides K4+5 (Table I) and K4+5' (data not shown) were similar. Automated Edman degradation of peptide K4+5 revealed a 29-residue sequence in which the second Lys residue was placed at the 10th cycle and two Met residues were placed at the 6th and 13th cycles (Fig. 1). Peptides K2 and K3 yielded 18- and 10-residue sequences, respectively.

Tryptic Digestion—Seven peptides, T1, T2, T2', T3+4+5, T6+7, T8+9, and T8+9' generated from PE-GFMEP (2 nmol) by tryptic digestion were separated on an Aquapore RP-300 column, as shown in Fig. 2B. Tryptic digestion was incomplete as evidenced by the presence of undigested protein peak eluted following peptide T1. As seen in Table I, amino acid compositions of T2 and T2' were identical. Those of T8+9 and T8+9' were very similar (data not shown) not only to each other but also to those of K4+5 (Table I) and K4+5'. Amino acid compositions of T3+4+5 (Table I) and T6+7 (data not shown) revealed that they were overlap peptides containing some Arg and/or Lys residues unsusceptible to tryptic digestion even in the presence of 2 M urea. Four peptides, T1, T2, T2', and T3+4+5 were subjected to amino-terminal sequence analyses (Fig. 1). Peptide T1 turned out to be the amino-terminal fragment as expected from its high content of Ala. Although peptide T2 did not yield any significant PTH signals in the second cycle of the Edman degradation, T2' yielded PTH-Thr. T3+4+5 was sequenced up to 40 cycles, although some residues remained unidentified.

CNBr Cleavage—Fragments of PE-GFMEP (5 nmol) by CNBr cleavage were first fractionated by GPC-HPLC as shown in Fig. 2C and four peptides were isolated by RP-HPLC on an Aquapore RP-300 column. Fragments M2+3 and M2+3+4 are large peptides containing one and two uncleaved Met residues, respectively. The sequence analysis together with MALDI-TOF MS analysis of the carboxyl-terminal fragment M4 (devoid of homoserine) confirmed the carboxyl terminus of GFMEP (Fig. 1).

Subdigestion with S. aureus V8 Protease—To obtain the sequence of the missing portions and to overlap fragmental sequences, fragments M1 and M2+3 were digested with S. aureus V8 protease. Two peptides each were separated from the digests by RP-HPLC on an Aquapore RP-300 column, M1-E1 and M1-E2 from M1 and M2+3-E1 and M2+3-E2 from M2+3. Analysis of peptide M1-E2 completed the sequence of peptide T1 and overlapped to peptide T2 and to T3+4+5 (Fig. 1). The sequence of peptide M2+3-E1 filled the gaps left in that of peptide T3+4+5 and extended the continuous amino-terminal sequence to residue 91. Likewise, the sequence of M2+3-E2 provided overlaps in the carboxyl-terminal portion resulting in the alignment of K peptides in the order of K2-K3-K4+5.

BNPS-Skatole Cleavage—Cleavage of tryptophyl bonds in PE-GFMEP (5 nmol) by BNPS-skatole resulted in the formation of five fragments, W1, W2, W3, W1+2, and W2+3, the latter two being overlap fragments. They were isolated by GPC-HPLC on a TSKgel G2000SWxL column followed by RP-HPLC on an Aquapore RP-300 column (data not shown). The 25-residue sequence of a lysine-containing carboxyl-terminal fragment W3 confirmed the overlap between K2 and M2+3-E2 (Fig. 1).

Cleavage with Endoproteinase Asp-N—To fill the gap left between the amino- and carboxyl-terminal sequences, endoproteinase Asp-N was employed to generate a peptide D4. The digest of PE-GFMEP (5 nmol) was separated by RP-HPLC on
an Aquapore RP-300 column (data not shown). Among the peptides isolated, a peptide D4 was selected on the basis of PE-Cys- and Trp-containing (254 and 290 nm-absorbing) peptide and subjected to Edman degradation. The 34-residue sequence obtained provided a solid overlap between T3 4+5 and K2 and resulted in the completion of the 167-residue sequence of GFMEP.

**Assignment of Disulfide Bonds**—To assign the locations of the two disulfide bonds, the CNBr-cleaved native GFMEP (5 nmol) was digested with trypsin and chymotrypsin and subjected to RP-HPLC. A peptide fraction that was not present in the reduced digest was applied to automated Edman degradation with 100 °C with 4 M trifluoroacetic acid for 3 h, and the sugar released was pyridyldiaminated and analyzed by HPLC. A TSKgel Sugar AXI column using pyridylaminated derivatives of glucose, mannose, and galactose as the standards (16). A single peak at retention time corresponding to pyridyldiaminated mannose was detected (data not shown). Mannosyl residue was completely split off from T2 by the action of jack bean α-mannosidase, and the resulting T2 yielded a peak product on an Aquapore RP-300 column at the same retention time as that of T2′ (data not shown). The susceptibility of mannose residue in T2 to jack bean α-mannosidase presented an unambiguous proof for α-configuration of glycosidic linkage to Thr42 in GFMEP.

**Sequence Analysis of POMEP**—Compositional analysis of POMEP revealed that this protein completely lacks two valuable residues, Met and Lys, for sequence analysis (Table II). Resistance of the protein, especially in the amino-terminal half, even after reduction and pyridylethylation in complete denaturing condition (6 M guanidine HCl), to proteolytic digestions with a prior denaturing step at 8 M urea combined with the absence of the strategic residues required use of a rather uncommon fragmentation procedure: tryptic digest following chemical cleavage by cyanolation (12). Fig. 4 shows the complete amino acid sequence of POMEP derived from the sequence data of tryptic fragments of PE- and S-cyanlated POMEP as well as those of BNPS-skatole cleavage of PE-POMEP. Overlaps between the major fragments were obtained from chymotryptic peptide W1-CH1 of a BNPS-skatole fragment W1, tryptic peptide TD1 of PE-POMEP pre-treated with 2% (v/v) formic acid, thermolytic peptides TH1, TH2, and TH3, peptide TRP-1 generated from PE-POMEP by successive digestion with trypsin and GFMEP, and peptide T7-D3 from endoproteinase Asp-N digestion of T7, and peptide E4 from V8 protease digestion of PE-POMEP. Edman degradation of the native protein revealed the amino-terminal sequence of 36 residues, except for the 6th and 34th residues where no significant PTH-derivative was observed.

**Tryptic Digestions**—Fig. 5A shows the elution profile of six peptides, T1, T2 +3+4, T5, T5 +6, T6, and T7, which were generated from PE-POMEP (1 nmol) by tryptic cleavage. Table II summarizes amino acid compositions and molecular mass values of PE-POMEP and these peptides. The largest tryptic peptide T2 +3+4 contained three Arg residues. Complete sequences of T1, T5, and T6 were obtained by Edman degradations, which were in accord with their amino acid compositions and mass values. As seen in Fig. 4, peptide T1 confirmed the amino-terminal 10-residue sequence and placed a Cys as the 6th residue. T7 must be the carboxyl-terminal fragment, which is the only peptide devoid of arginine. The largest peptide, T2 +3+4, was refractory to Edman degradation. Because T5, T6, and T7 do not have the amino-terminal sequence of

---

**Amino Acid Sequences of Metalloendopeptidases**

**TABLE I**

| Amino acid compositions of PE-GFMEP and its peptides from GFMEP or tryptic digest |
| PE-GFMEP (167–198) | T1 (1–40) | T2 (41–52) | T3 (53–124) |
|---------------------|---------|---------|---------|
| Asp/Asn (D/N) | 17.0 (8/9) | 5.9 (3/4) | 20.0 (20/20) | 20.0 (20/20) |
| Glu/Gln (E/Q) | 12.7 (5/7) | 6.1 (2/4) | 23.1 (23/23) | 23.1 (23/23) |
| Ser (S) | 19.2 (22) | 11.8 (16) | 26.3 (3) | 26.3 (3) |
| Gly (G) | 3.9 (10) | 6.1 (5) | 33.4 (9) | 33.4 (9) |
| His (H) | 5.7 (6) | 3.0 (3) | 16.2 (1) | 16.2 (1) |
| Arg (R) | 3.8 (4) | 3.2 (3) | 0.9 (1) | 0.9 (1) |
| Thr (T) | 15.1 (19) | 11.1 (12) | 4.6 (6) | 4.6 (6) |
| Ala (A) | 24.0 (24) | 13.0 (14) | 26.2 (2) | 26.2 (2) |
| Pro (P) | 5.5 (5) | 2.5 (2) | 1.1 (1) | 1.1 (1) |
| Tyr (Y) | 12.5 (13) | 8.3 (10) | 4.6 (6) | 4.6 (6) |
| Val (V) | 6.0 (6) | 4.0 (4) | 1.0 (1) | 1.0 (1) |
| Met (M) | 3.1 (3) | 0.7 (1) | 0.9 (2) | 0.9 (2) |
| PE-Cys (C) | 4.4 (4) | 3.0 (4) | 1.1 (1) | 1.1 (1) |
| Ile (I) | 0.9 (1) | 0.8 (1) | 1.0 (1) | 1.0 (1) |
| Leu (L) | 7.4 (7) | 4.7 (5) | 1.0 (1) | 1.0 (1) |
| Phe (F) | 8.1 (8) | 5.8 (6) | 1.0 (1) | 1.0 (1) |
| Trp (W) | 2.9 (3) | 1.2 (1) | 1.2 (1) | 1.2 (1) |
| Lys (K) | 4.2 (4) | 0.8 (1) | 1.0 (1) | 1.0 (1) |

**Yield (%)** 43.7 27.6 59.3 5.5 28.4 17.5 8.2 24.3

---

*a Each composition was calculated on the basis of the integral value of the residue marked.

*b The symbol + indicates the presence of absorbance at 290 nm.
QTQLN– (starting at residue 11), it is likely that the amino-terminal Gln in the peptide T2+3+4 was converted to a pyroglutamyl residue. To obtain sequence information in the amino-terminal half of the molecule, POMEP was cleaved chemically by cyanylation and digested with trypsin. Of isolated fragments, TCN3, TCN4, TCN5, TCN6+7, and TCN8 were unique to CN-POMEP (data not shown) and subjected to Edman degradations (Fig. 4). The 25-residue sequence of TCN3...
confirmed the amino-terminal sequence analysis results of the intact protein. Fragments TCN4 and TCN5 yielded 12- and 19-residue sequences, respectively, whereas TCN6 and TCN7 were found to be amino-terminally blocked. To fill the big hole left in the middle of the molecule, PE-POMEP was treated with 2% (v/v) formic acid and then digested with trypsin. A Trp- and PE-Cys-containing peptide TD1 was isolated from the digest by RP-HPLC and subjected to Edman degradation. The 20-residue sequence of TD1 was extended by the analysis of another Trp-containing peptide TKAE-1 isolated from a consecutive digest of AE-POME P with trypsin and GFMEP.

**BNPS-Skatole Cleavage**—Cleavage of PE-POME P by BNPS-skatole yielded four distinctive peaks, W1, W1+2, W2+3, and W3, by GPC-HPLC on tandem columns of TSKgel G2000SWXL and TSKgel G3000SWXL (Fig. 5B). As seen in Fig. 4, the sequence of W2+3 provided an overlap between TCN4 and TCN5 and that of W3, TKAE-1 and T5. The sequence of the amino-terminal half was completed by the analysis of the chymotryptic peptide, W1-CH1, of fragment W1, although overlap between W1-CH1 and TCN4 was marginal.

**Carboxyl-terminal Sequence**—Peptide T7-D3 isolated from the endoproteinase Asp-N digest of T7 extended the carboxyl-terminal sequence. The amino acid composition of T7 along with the observed sequence and mass of T7-D3 (MH⁺: 1,576) indicated that the peptide is composed of 14 residues, of which the two unidentified residues in the carboxyl terminus must be filled with Ala and Leu. Capillary liquid chromatography/FAB MS analysis of *S. aureus* V8 protease digest of PE-POME P yielded a clear sequence of NNPALA among several sequences (data not shown). This sequence established the carboxyl terminus of the protein.

**Remaining Overlaps and the Central Gap**—It remained to overlap tryptic peptides T5, T6, and T7 and to fill the gap left between TCN5 and TD1. Thermolytic peptides of PE-POME P provided this information. Arginine-containing peptides TH2 and TH3 overlapped T5 with T6 and T6 with T7, respectively. The 11-residue sequence of a PE-Cys-containing peptide TH1 extended the sequence of TD1 amino-terminally but was too short to overlap to observed sequence of TCN5. The calculated 2,142.3 atomic mass units for the 19-residue sequence of TCN5 was not in accord with the observed mass value of this peptide (MH⁺: 2,505) by MALDI-TOF MS. This discrepancy of 363 atomic mass units indicated the presence of unidentified residues at the carboxyl terminus of TCN5. The peptide TCN5 was generated by the cleavage of acyl-Cys bond followed by tryptic digestion. All four Cys as well as six Arg residues were placed firmly in the sequence. Thus the unidentified carboxyl-terminal sequence of TCN5 must be Thr-Phe-Asp with the additional mass value of 363.4 atomic mass units, of which Phe-Asp is the amino terminus of TH1. The molecular weight of POME P determined by MALDI-TOF MS was 17,926, which is in agreement with the calculated value of 17,921.2. This result supports the hypothesis that the sequence of 168 residues is valid.

**Hydropathy Profile, Secondary Structure Predictions, and Charge Distribution**—Fig. 6 depicts the hydropathy profiles,

---

**Fig. 5.** Fractionation of peptides of PE-GFMEP. A, tryptic peptides were separated by RP-HPLC on an Aquapore RP-300 column (2.1 x 30 mm) with a linear gradient of acetonitrile (0–64%) in 16 min at a flow rate of 0.2 mL/min. B, products of the chemical cleavage at tryptophanyl residues with BNPS-skatol were separated by GPC-HPLC on tandem columns (7.8 x 300 mm, each) of TSKgel G2000SWXL and TSKgel G3000SWXL.

**Fig. 6.** Comparisons of hydropathy profile, predicted secondary structure, and charge distribution of GFMEP and POME P. The hydropathy profiles (upper panels) of GFMEP (panel A) and POME P (panel B) were calculated by the method of Kyte and Doolittle (33) with a window size of 7, respectively. The secondary structure prediction (middle panel) was carried out by the method of Chou and Fasman (34). Hatched boxes, white boxes, and lines denote α-helix, β-sheet, and extended structure, respectively. The locations of positive (Arg, Lys, and His) and negative (Asp and Glu) charges in the polypeptides are indicated by upper and lower bars, respectively.
predicted secondary structures, and charge distributions of GFMEP and POME. From the data, it is obvious that both MEps are relatively hydrophilic proteins. The sequence motifs, HEXXX(His)_{11-13}-His^{121} for GFMEP and His^{119}, His^{122} for POME, in the hydrophilic region of the carboxy-terminal portion of both enzymes are the putative zinc-binding sites by an analogy with the primary structures of various metalloproteinases. Interestingly, it was found that these sequence motifs were predicted to be in the random coil regions.

DISCUSSION

GFMEP is composed of a single polypeptide chain of 167 amino acid residues with two intra-chain disulfide bonds (Cys^{6} to Cys^{76} and Cys^{77} to Cys^{97}), as shown in Fig. 1. Although in less significant amounts, PTH-cystine was also observed at the 14th cycle in the Edman degradation of the disulfide-bridged CNBr fragment. Thus the assignment of disulfide bonds should be regarded as tentative. A single mannose residue is attached to Thr^{42} as an unusual post-translational modification. The resistance of the bonds should reflect the substrate-binding site structure of the enzyme. A kinetic study using synthetic peptide substrates is in progress.

POME is homologous to GFMEP as shown in Fig. 7 with a single residue extension at the amino terminus. Two intrachain disulfide bonds are assumed to be conserved (Cys^{6} to Cys^{76} and Cys^{77} to Cys^{97}). The molecular weight is calculated to be 17,921.2, which is in accord with the observed weight (M+H)^{+}) of 17,927 determined by MALDI-TOF MS. Although they share 61% sequence identity, analysis of POME was hampered by the absence of two key residues, Met and Lys, and by the resistance to proteolytic digestions in the amino-terminal half of the molecule.

Both GFMEP and POME were found to share significant sequence homology (21–26% identity) with a family of metalloproteinases categorized as deuterolysins in a search for homologous proteins (Fig. 7) with the BLAST program (17). They are Aspergillus oryzae neutral protease II (accession number 1078624) (18), Aspergillus flavus 24-KDa metalloproteinase precursor (deuterolysin, accession number 1170905), Aspergillus fumigatus (listed as Sartorya fumigata) in protein sequence data base MEP20, (accession number 780794) (19), and penicilloly-XX (listed as Sartorya fumigata) in protein sequence data base MEP20, (accession number 780794) (19), and penicillolysis needle protease II (accession number 1078641) from Penicillium citrinum (20). From the sequence alignment, it is apparent that they have evolved from a common ancestor, although only mushroom enzymes have the unique specificity. The consensus sequence, HEXXXH, in these MEps corresponds to the zinc-binding sequence motif, HEXXH, in thermolysin and related zinc metalloproteinases. A zinc atom in thermolysin has been demonstrated to be bound to His^{142}, His^{146}, and Glu^{166} by extensive studies of the tertiary structure (21). Bode et al. (22), working with crystals of asatin, a metalloproteinase from crayfish, have identified His^{52}, His^{98}, His^{102}, and Tyr^{149} as the ligands of zinc, which are distinct from those of thermolysin and other related proteases. On the basis of the sequence homology for the zinc-binding regions, metalloproteinases are classified into several subfamilies (8, 23). We therefore compared sequences around the putative zinc-binding sites of GFMEP and POME as well as those of deuterolysins with the zinc-chelating sequences in typical metalloproteinases. As shown in Fig. 8, there was no homology among these families with the exception of the sequence of HEXXXH. However, the figure also indicated the presence of two homologous regions, GTXDXXYG and AXXX-NXD, among the sequences of six fungus metalloproteinases, in which potential zinc ligand residues Asp, Asn, and Tyr were conserved. These observations suggested that Asp, Asn, and/or Tyr in these regions are likely the third and/or fourth zinc ligand in these metalloproteinases and that the six fungus metalloproteinases might be classified into a novel subfamily of zinc metalloproteinases. Efforts to elucidate the three-dimensional structure of GFMEP by x-ray crystallography are now underway.

Mass spectral analyses of two tryptic peptides T2 and T2' of GFMEP with identical amino acid compositions presented a clue for the partial α-mannosylation at Thr^{42}. Attachment of a single sugar residue to the hydroxy group of Ser or Thr residue has received particular attention, because O-GlcNAc bound to the hydroxy amino acid residues has been extensively investigated with respect to the functions of the proteins (24). Recently, an L-fucosyl residue bound to Thr^{42} of O-glycosidically has been characterized in several proteins (25–28). However, to our best knowledge, identification of α-mannosylated GFMEP is the first report that presented unambiguous evidence for the O-glycosidic linkage of a D-mannose residue to the hydroxy group of a Thr residue in protein. On the other hand, attach-

---

8 N. Dohmae, T. Nonaka, Y. Hashimoto, and K. Takio, unpublished observation.
Fig. 8. Comparison of the sequences around the putative zinc-binding sites of zinc metalloproteinases. Panels A and B show consensus sequences of representative metalloproteinase subfamilies, gluzincins and metzincins, respectively. Asterisks indicate zinc ligands determined by x-ray crystallography. Sequences around HEXXH of six fungus metalloproteinases are aligned in panel C, and putative zinc ligands are indicated by +. Common residues are indicated by white letters on a black background.

Acknowledgment—We thank Dr. Yoichi Tsunuraya for valuable suggestions for sugar analysis.

Note Added in Proof—After acceptance for publication, a homologous sequence was found in the GenBank™ database. Extracellular protease of Aeromonas hydrophila (accession 1945442) shares sequence identity of 44.6 or 44% with GFMEP or POMEP, respectively, with most of the conserved residues in Fig. 7 conserved. The sequence was submitted to GenBank™ directly, and detailed description of the enzyme (including whether the protein is actual one or simply a probable protein translated from an open reading frame in genomic DNA) has not been published. Recent preparations of GFMEP (preparared for crystal) showed complex mass spectral pattern indicating multiple hexose attachment (up to 4 hexose units). Those are likely bound to the same Thr because RP-HPLC profile around the elution position of trypptic peptide T2 (Fig. 2B) became complex, and pooled fractions gave similar multiple hexose attachment signals in MALDI-TOF MS spectra. It is not known why the hexose content would be different from that reported here; either strain or culture conditions might have been changed by the grower.

REFERENCES
1. Terashita, T., Oda, K., Kono, M., and Murao, S. (1985) Trans. Mycol. Soc. Jpn. 26, 397–409
2. Terashita, T., Oda, K., Kono, M., and Murao, S. (1985) Agric. Biol. Chem. 49, 2933–2930
3. Lewis, W. G., Basford, J. M., and Walton, P. L. (1978) Biochim. Biophys. Acta 522, 551–560
4. Shipolini, R. A., Callewaert, G. L., Cottrell, R. C., and Vernon, C. A. (1974) Eur. J. Biochem. 48, 465–476
5. Doonan, S., Doonan, H. J., Hanford, R., Vernon, C. A., Walker, J. M., Airedil, L. P. da S., Bossa, F., Barra, D., Carlion, M., Fasella, P., and Riva, F. (1975) Biochem. J. 149, 497–506
6. Nonaka, T., Ishihikawa, H., Tsunuraya, Y., Hashimoto, Y., Doehme, N., and Takio, K. (1993) J. Biochem. (Tokyo) 118, 1014–1020
7. Doehme, N., Hayashi, K., Miki, K., Tsunuraya, Y., and Hashimoto, Y. (1995) Biosci. Biotechnol. Biochem. 59, 2047–2050
8. Hopper, N. H. (1994) FEBS Lett. 354, 1–6
9. Cole, R. D. (1967) Methods Enzymol. 11, 315–317
10. Gross, E. (1967) Methods Enzymol. 11, 238–255
11. Omenn, G. S., Fontana, A., and Anfinsen, C. B. (1970) J. Biol. Chem. 245, 1895–1902
12. Stark, G. R. (1977) Methods Enzymol. 47, 129–132
13. Inglis, A. S. (1983) Methods Enzymol. 91, 324–332
14. Billingmeyer, B. A., Cohen, S. A., and Tarvin, T. L. (1984) J. Chromatogr. 336, 93–104
15. Hewick, R. M., Hunkapiller, M. W., Hoed, L. E., and Dreyer, W. J. (1981) J. Biol. Chem. 256, 7990–7997
16. Suzuki, K., Jondo, A., Kato, I., Hase, S., and Ikennaka, T. (1991) Agric. Biol. Chem. 55, 283–284
17. Altcheul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
18. Tatsusi, H., Murakami, S., Tsuji, R. F., Ishida, Y., Murakami, K., Masaki, A., Kawabe, H., Arimura, H., Nakano, E., and Motai, H. (1991) Mol. Gen. Genet. 228, 97–103
19. Ramesh, M. V., Sirakova, T. D., and Kolattukudy, P. E. (1995) Gene (Amst.) 165, 121–125
20. Matsumoto, K., Yamaguchi, M., and Ichishima, E. (1994) Biochim. Biophys. Acta 1218, 469–472
21. Mathews, B. W., Janssenius, J. N., Coleman, P. M., Schoenborn, B. P., and Dupourque, D. (1972) Nat. New Biol. 238, 37–41
22. Bode, W., Gomis-Riu, F.-X., Huber, R., Zwilling, R., and Stöcker, W. (1992) Nature 358, 164–167
23. Stücker, W., Grana, F., Baumann, U., Reinemer, P., Gomis-Riu, F.-X., Mckay, D. B., and Bode, W. (1995) Protein Sci. 4, 823–840
24. Haitwanger, R. S., Holt, G. D., and Hart, G. W. (1990) J. Biol. Chem. 265, 2563–2568
25. Kentzer, K. J., Buko, A., Menon, G., and Sarin, V. K. (1990) Biochem. Biophys. Res. Commun. 171, 401–406
26. Harris, R. J., Leonard, C. R., Guzzetta, A. W., and Spellman, M. W. (1991) Biochemistry 30, 2311–2314
27. Harris, R. J., Ling, V. T., and Spellman, M. W. (1992) J. Biol. Chem. 267, 5102–5107
28. Stults, N. L., and Cummings, R. D. (1993) Glycobiology 3, 589–596
29. Stahl-Bolanger, S., and Tanner, W. (1991) Eur. J. Biochem. 196, 185–190
30. Sprio, R. G., and Bhoyroo, V. D. (1980) J. Biol. Chem. 255, 5347–5354
31. Yun, S.-J., Hirooka, Y., Nishizawa, M., Takio, K., Titani, K., Nogi, Y., and Fukasawa, T. (1991) J. Biol. Chem. 266, 693–697
32. Doonan, S., and Fahmy, H. M. A. (1975) Eur. J. Biochem. 56, 421–426
33. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
34. Chou, P. Y., and Fasman, G. D. (1978) Annu. Rev. Biochem. 47, 251–276
35. Higgins, D. G., and Sharp, P. M. (1988) Gene (Amst.) 73, 237–244