Note

Complete Amino Acid Sequence of Chitinase-A from Leaves of Pokeweed (Phytolacca americana)

Takeshi YAMAGAMI,† Miho TANIGAWA, Masatsune ISHIGURO, and Gunki FUNATSU

Laboratory of Protein Chemistry and Engineering, Faculty of Agriculture, Kyushu University, Fukuoka 812-8581, Japan

Received November 18, 1997

The complete amino acid sequence of pokeweed leaf chitinase-A was determined. First all 11 tryptic peptides from the reduced and S-carboxymethylated form of the enzyme were sequenced. Then the same form of the enzyme was cleaved with cyanogen bromide, giving three fragments. The fragments were digested with chymotrypsin or Staphylococcus aureus V8 protease. Last, the 11 tryptic peptides were put in order. Of seven cysteine residues, six were linked by disulfide bonds (between Cys25 and Cys74, Cys89 and Cys98, and Cys195 and Cys208); Cys176 was free. The enzyme consisted of 208 amino acid residues and had a molecular weight of 22,391. It consisted of only one polypeptide chain without a chitin-binding domain. The length of the chain was almost the same as that of the catalytic domains of class II chitinases. These findings suggested that this enzyme is a new kind of class II chitinase, although its sequence resembles that of catalytic domains of class II chitinases more than that of the class III chitinases reported so far. Discussion on the involvement of specific tryptophan residue in the active site of PLC-A is also given based on the sequence similarity with rye seed chitinase-

Key words: chitinase; amino acid sequence; Phytolacca americana; pokeweed

Plant chitinases (EC 3.2.1.14) catalyze the hydrolysis of β-1,4-linked homopolymers or oligomers of N-acetylglucosamine, can be classified roughly into three classes:1-23 class I, which has catalytic (Cat) and chitin-binding (CB) domains; class II, with only a Cat domain, which resembles that of class I; and class III, with no sequence similarity to class I and II chitinases. Cat domains, which participate in the hydrolysis of the substrates of class I and II chitinases, are further classified into two subclasses depending on their high or low molecular mass.23

In a series of studies on the structure-function relationships of plant chitinases, we isolated and characterized class IH and class IIH chitinases, referred to here as RSC-a and RSC-c, respectively, from rye (Secale cereale) seeds,9 and also class II and class III chitinases, referred to here as PLC-A and PLC-B, respectively, from pokeweed (Phytolacca americana) leaves,9 and we sequenced the amino acid residues of RSC-a,9 RSC-c, and PLC-B.23 In this paper, we report the complete amino acid sequence of PLC-A and compare its structure and function with those of other chitinases.

PLC-A was prepared from pokeweed leaves as described before.23 The amino acid composition of PLC-A, derived from protein hydrolysate was Asp32.8, Thr10.7, Ser15.9, Glu16.6, Pro10.9, Gly23.7, Ala21.7, Val11.8, Met0.6, Ile8.0, Leu8.1, Tyr9.1, Phe14.0, Lys6.0, His4.1, Arg4.8, Trp4.0, and 1/2Cys5.4 moles per mole of PLC-A. Examination of Rcm-PLC-A with a gas-phase sequencer (Shimadzu PSQ-1) gave the sequence up to residue 26 (Fig. 1). Carboxypeptidase Y was used to digest 1 nmol of Rcm-PLC-A and released 0.6 nmol of CM-cysteine and 0.1 nmol of threonine by 10 min; 0.9 nmol of CM-cysteine, 0.8 nmol of threonine, and 0.6 nmol of leucine by 30 min; and 1.0 nmol of CM-cysteine, 0.9 nmol of threonine, and 0.6 nmol of leucine by 60 min, respectively, indicating the C-terminal sequence of PLC-A to be -Leu-Thr-Cys.

Rcm-PLC-A (3 mg) prepared by the method of Crestfield et al.,9 was digested with 1/100 (w/w) TPCK-trypsin (Sigma Chemical Co.) in 0.2 M ammonium acetate buffer, pH 7.8, at 37°C for 3 h and lyophilized. A tryptic digest of Rcm-PLC-A was dissolved in 5 mM potassium phosphate buffer, pH 6.0, and separated into soluble (T₁) and pellet (T₂) fractions by centrifugation. Peptides in the T₁-fraction were separated by reverse-phase HPLC on a C4 column (4.6 × 250 mm) with a phosphate-McCN system,9 yielding six peptides (T₂, T₃, T₆, T₉, T₁₀, and T₁₁) and free lysine (numbered T₅). Each peptide was further purified by reverse-phase HPLC on a C4 column (4.6 × 150 mm) with a TFA-MeCN system.9 Peptides in the T₂-fraction were first separated by reverse-phase HPLC on a C4 column (4.6 × 150 mm) with a TFA-MeCN system, yielding four peptides (T₁, T₄, T₇, and T₈), and then purified by the same column with an ammonium acetate-MeCN system.9 Thus, all 10 tryptic peptides and also free lysine, numbered T₁-T₁₁, were isolated from the tryptic digest of Rcm-PLC-A. The amino acid composition was analyzed by a Hitachi 655-A or Pico Tag amino acid analyzer. Sequencing of these peptides except for peptide T₄ were done directly by the DABITC-PITC double coupling method10 or by a gas-phase amino acid sequen-
Peptide T4 was sequenced up to residue 18.

Peptide T4 was digested with 1/20 (w/w) *Staphylococcus aureus* V8 protease (Miles Co.) in 50 mM ammonium hydrogen carbonate buffer, pH 7.8, containing 2 mM EDTA and 4 M urea at 37°C for 5 h. The resulting peptides, T4V1 and T4V2, were separated by reverse-phase HPLC on a C8 column (4.6 x 150 mm) with a TFA-MeCN system. Peptide T4V2 was sequenced up to residue 24.

RCm-PLC-A was cleaved with cyanogen bromide by the method of Steers et al. and the resulting three fragments, CBI, CBI', and CBII, were purified by gel filtration on a Bio-Gel P-30 column in 30% aqueous acetic acid and then reverse-phase HPLC on a C4 column.

RCm-CBI was digested with 1/50 (w/w) TLCK-chymotrypsin (Sigma) in 0.2 M N-methylmorpholine buffer, pH 8.1, at 37°C for 3 h and the resulting peptides were separated by reverse-phase HPLC on a C4 column (4.6 x 250 mm) with a phosphate-MeCN system, yielding 12 peptides, CBI-C1 to -C12. Sequencing of peptide CBI-C6 made clear the sequence of peptide T4. The sequencing of peptides CBI-C1, -C2, -C3, -C6, and -C7 made possible the alignment of T1 with T2, T2 with T3, T3 with T4, T4 with T5 and T6, and T6 with T7, respectively.

RCm-CBI' was digested with 1/15 (w/w) *S. aureus* V8 protease in 50 mM sodium hydrogen carbonate buffer, pH 7.8, containing 2 mM EDTA and 4 M urea at 37°C for 12 h and the resulting peptides were separated by reverse-phase HPLC on a C4 column (4.6 x 150 mm) with a TFA-MeCN system, yielding two peptides, CBI'-V1 and CBI'-V2. The sequencing of peptide CBI'-V2

---

**Fig. 1.** Complete Amino Acid Sequence of Pokeweed Leaf Chitinase-A.

T and CB under the sequence indicate peptides obtained by tryptic hydrolysis and cleavage by CNBr of RCm-PLC-A, respectively. I-C indicates peptides obtained by chymotryptic hydrolysis of CNBr fragment CBI from RCm-PLC-A. 1'V and II-V indicate the peptides derived from *S. aureus* V8 protease digestion of CNBr fragments CBI and CBII from RCm-PLC-A, respectively. Individual peptides are described as follows: ----, sequenced with DABITC-PITC method or gas-phase sequencer; ----, not identified by sequencing. The arrow indicates the amino acid residues identified by direct sequencing of PLC-A.
made possible the alignment of T7 with T8. RCm-CBII was digested with 1/30 (w/w) *S. aureus* V8 protease in 50 mM sodium hydrogen carbonate buffer, pH 7.8, containing 2 mM EDTA and 4 mM urea at 37°C for 12 h and the resulting peptides were separated by reverse-phase HPLC on a C4 column (4.6 × 150 mm) with a TFA-MeCN system, yielding five peptides, CBII-V1 to -V5. The sequencing of peptides CBII-V1 and CBII-V3 made possible the alignment of T8 with T9 and T9 with T10 and T11, respectively.

Measurement of free sulphydryl groups of PLC-A by the method of Habeeb[12] showed that PLC-A contained 0.85 mol of sulphydryl groups per mole of PLC-A. From this result, we concluded that six cysteine residues were linked by disulfide bonds and that one cysteine residue was free. To identify which cysteine residues were disulfide-bridged and which residue was free, we treated S-carboxymethyl- (CM-) PLC-A with CNBr and filtered the reaction mixture on a Bio-Gel P-30 gel column, yielding two fragments, CB1 and 2. Fragment CB1, which contained Cys25, 74, 89, and 98, was digested first with 1/50 (w/w) lysylendopeptidase (Wako Pure Chemical Industries) in 50 mM MES buffer, pH 6.5, containing 4 mM urea at 37°C for 2 h. The digest was diluted twice with 50 mM MES buffer, pH 6.5, and further digested with 1/30 (w/w) *S. aureus* V8 protease for 4 h and then with 1/600 (w/w) endoproteinase Asp-N (Boehringer Mannheim Industries) for 18 h. The resulting peptides were separated by reverse-phase HPLC on a C4 column (4.6 × 250 mm) with a phosphate-MeCN system, yielding two peptides containing disulfide, CB1-SS-1 and -2. From the amino acid sequences, Cys25 and Cys74 were found to be linked by disulfide bonds, as were Cys89 and Cys98 (Fig. 2). Fragment CB2, which contained Cys176, 195, and 208, was digested first with 1/50 (w/w) TPCK-trypsin for 2 h, then with 1/50 (w/w) TLCK-chymotrypsin for 2 h in 50 mM sodium phosphate buffer, pH 6.5, containing 2 mM urea. The resulting peptides were separated by reverse-phase HPLC on a C4 column (4.6 × 250 mm) with a phosphate-MeCN system, yielding two peptides containing disulfide, CB2-SS-1 and -2, and three peptides containing CM-cysteine, CB2-S-1 to -3. Their amino acid sequences are shown in Fig. 2. Peptides CB2-S-1 and -2 were found to be derived by cleavages of Arg-Ile and Phe-Tyr bonds, respectively, in peptide CB2-SS-2. These results showed that Cys195 and Cys208 were linked by a disulfide bond and that the sulphydryl group of Cys 176 was free.

The amino acid composition of PLC-A, calculated from its sequence, was Asp14, Asn19, Thr11, Ser15, Glu7, Gln10, Pro9, Gly23, Ala21, Val12, Met1, Ile9, Leu7, Tyr12, Phe14, Lys6, His4, Arg5, Trp3, and Cys7, which agrees with that found by chemical analysis of PLC-A. PLC-A consisted of 208 amino acid residues and its molecular weight was calculated to be 22,391, the smallest for plant chitinases reported so far. There being 15 basic and 21 acidic residues was compatible with the isoelectric point of the protein being 3.7, as was found experimentally.

The amino acid sequence of PLC-A is compared with sequences of the Cat domain of chitinases of class IH (RSC-a), class II (ynam),[13] class IIL (tomato),[14] and class III (tomato) in Fig. 3. Class IIL chitinases are composed of 243–245 residues, so PLC-A (208 residues), lacking a chitin-binding domain, is a class IIL chitinase. The chain length and C-terminal sequence of PLC-A ending with a cysteine residue are, however, similar to those of the Cat domain of class IIL chitinases from yam (209 residues),[13] bean (207 residues),[13] corn (205 residues),[14] and rapeseed (207 residues),[17] but they are not similar to those of the class IIL chitinases reported so far from tobacco (229 residues),[13] petunia (231 residues),[13] and tomato (231 residues).[14] PLC-A had 59–64% sequence similarity to class IIL enzymes, 42–45% similarity to class ILM enzymes, and 47–51% similarity to class IH and IIL enzymes. Seven cysteine residues in PLC-A were conserved in all these class IIL, II, and IIL chitinases, but chitinases classified so far as being in class IIL lack the disulfide bridge between Cys89 and Cys98 found in PLC-A. These results suggested that PLC-A is a new kind of class IIL chitinase, and that the chitinases from tobacco, petunia, and tomato can be grouped as another kind of class IIL chitinase.

In a previous paper,[19] we showed that Trp72 in RSC-c is important for enzyme activity. Findings about the three-dimensional structure of barley seed chitinase,[20] which has 95% sequence similarity to RSC-c, suggest that Trp72 of RSC-c is at the far end of the active site cleft (probably at the substrate-binding site) and may affect binding of low-molecular-weight substrates such as (GlcNAc)₆.

For the optimal alignment of PLC-A with RSC-c, shown in Fig. 3, three deletions (of residues 70–82, 162–167, and 185–197 in RSC-c) were necessary. The seven cysteine residues in the Cat domains of class IH, II, and IIL chitinases are all conserved in PLC-A and six are linked by disulfide bridges. PLC-A, in which an active-site tryptophan is deleted, may be a mutant of RSC-c. From contradictory experimental results,[42] that PLC-A easily hydrolyzes (GlcNAc)₂ into two moles of
Fig. 3. Amino Acid Sequence of PLC-A Compared with Sequences of Class II, III, IIA, and III Chitinases from Yam, Rye, Rye, and Tomato, respectively.

Amino acids identical to those in PLC-A are enclosed in boxes. Several gaps (−) have been inserted to give optimal alignment of the proteins.

(GlcNAC)2, and that RSC-c hydrolyzes (GlcNAC)2 little if at all, although both are active toward glycolchitin, a high-molecular-weight substrate, we infer that the deletion of the region containing Trp72 causes PLC-A to have high hydrolytic activity by removal of the interference of this region with the binding of (GlcNAC)2 to the substrate-binding site.

References