Amino Acid Sequence and Some Properties of Lectin-D from the Roots of Pokeweed (Phytolacca americana)

Ken-ichi Yamaguchi, Ayumi Mori, and Gunki Funatsu

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

Received February 21, 1996

Two pokeweed lectins, designated PL-D1 and PL-D2, have been isolated from the roots of pokeweed (Phytolacca americana) using chitin affinity column chromatography followed by gel filtration on a Sephacryl S-200 column and fast protein liquid chromatography on a Mono-Q column, and their amino acid sequences have been analyzed. PL-D1 consists of 84 amino acid residues and has a molecular mass of 9317, while PL-D2 has an identical sequence with PL-D1 except lack of the C-terminal Leu-Thr. PL-D is composed of two chitin-binding domains, A and B, with 50% homology with each other. Both PL-Ds did not agglutinate native rabbit erythrocytes, but showed about 0.1% of the agglutinating activity of wheat germ agglutinin toward trypsin-treated erythrocytes. In the presence of β(1→4) linked oligomers of N-acetyl-D-glucosamine, which inhibit the hemagglutination, PL-D1 had an ultraviolet-difference spectrum with maxima at 292–294 nm and 284–285 nm, attributed to the red shift of the tryptophan residue, suggesting the location of tryptophan residue(s) at or near saccharide-binding site of PL-D1.

Key words: lectin; chitin-binding protein; pokeweed; amino acid sequence

The roots of pokeweed (Phytolacca americana) contain several kinds of lectins specific for N-acetyl-D-glucosamine (GlcNAc) and some of them have mitogenic activity toward peripheral lymphocytes.1–31 We had isolated pokeweed lectin (PL)-A, PL-B, and PL-C from young pokeweed roots and shown that they are homologous proteins but have different molecular sizes and biological properties.4–6 Moreover, we have shown that PL-C, a dominant lectin in pokeweed roots, consists of three chitin-binding domains homologous to those of wheat germ agglutinin (WGA) and class I chitinases.5 We recently found that PL-A, PL-B, and PL-C are chitin-binding proteins and in addition there is another chitin-binding lectin, designated pokeweed lectin-D (PL-D), consisting of two chitin-binding domains in the pokeweed roots. In this paper, we describe the purification, amino acid sequence, and some properties of PL-D.

The crude lectin was prepared from the young roots of pokeweed as described previously.52 The crude lectin was dialyzed against 10 mm Tris-HCl buffer, pH 7.5, and put on a chitin column (2.8 x 18.5 cm), packed with chitin powder (Seikagaku Co.) and equilibrated with the same buffer. After washing out the non-adsorbed proteins with the same buffer, first the non-specifically adsorbed proteins were eluted with 1 M NaCl in the same buffer and then the remaining chitin-binding proteins with 0.5 M acetic acid solution. Gel filtration of the chitin-binding proteins on a Sephacryl S-200 column (2.0 x 140 cm) in 0.1 M Tris-HCl buffer, pH 7.5, gave two fractions: one contained PL-A, PL-B, and PL-C and the other a new pokeweed lectin designated PL-D. After dialysis against 10 mm sodium acetate buffer, pH 5.5, the PL-D fraction was put on a Mono-Q column (0.5 x 5.0 cm) in a fast protein liquid chromatography (FPLC) apparatus (Pharmacia Fine Chemicals, Uppsala, Sweden) and eluted by a linear gradient of NaCl from 0 to 0.5 M in the same buffer at a flow rate of 0.5 ml min

Fig. 1. FPLC-Elution Profile and Tricine SDS-Polyacrylamide Gel Electrophoresis of PL-D1 and PL-D2.

The PL-D fraction obtained by gel filtration on a Sephacryl S-200 column was put on a Mono-Q column (0.5 x 5 cm) equilibrated with 10 mm sodium acetate buffer, pH 5.5 and eluted by a linear gradient of NaCl from 0 to 0.5 M in the same buffer at a flow rate of 0.5 ml min. Fractions PL-D1 and PL-D2, indicated by the bar, were analyzed on a tricine SDS PAGE (inset). 1. PL-D1; 2. PL-D2; std. standard proteins: ovalbumin (45 kDa), z-chymotrypsinogen (24 kDa), myoglobin (17.2 kDa), cytochrome c (12.6 kDa), and aprotinin (6.5 kDa).
separating into two fractions (PL-D1 and PL-D2) at 0.32 and 0.34 M of NaCl (Fig. 1). Fractions PL-D1 and PL-D2, indicated by a bar, gave one band on SDS-PAGE, and their molecular masses were estimated to be approximately 9.5 and 9.4 kDa, respectively, by tricine SDS-PAGE in the presence of 2-mercaptoethanol (2-ME) by the method of Schägger and Jagow.98

The amino acid sequence of PL-D1 was analyzed by sequencing the peptides obtained by lysylendopeptidase, chymotryptic, and tryptic digestions as well as CNBr-cleavage of reduced and S-carboxymethylated (RPe)PL-D1 prepared by the method of Friedman et al. 79 Digestion of RPe-PL-D with lysylendopeptidase (Wako Pure Chem. Ind.) was done using 1/100 (w/w) enzyme in 4 M urea–50 mM Tris–HCl buffer, pH 9.0, at 37 C for 3 h. Fragmentation of the large peptide obtained was done using 1/100 (w/w) trypsin or 1/50 (w/w) chymotrypsin in 100 mM ammonium bicarbonate solution, pH 7.8, at 37 C for 3 h. The peptides were separated by reverse-phase (RP)-HPLC with a YMC-GEL C4 column (4.6 x 250 mm or 4.6 x 150 mm) by a linear gradient of acetonitrile from 0 to 30% in 0.1% trifluoroacetic acid solution. Amino acid analysis of peptides was done with a Pico-Tag amino acid analyzer (Waters Co.) after hydrolysis with constant-boiling HCl containing 0.05% 2-ME in vacuo at 110 C for 24 h. Amino acid sequence analysis was done by either manual Edman degradation using the DABITC/PTC double-coupling method 99 or an automated Shimadzu gas-phase protein sequencer PSQ-1.

The N-terminal sequences of RPe-PL-D1 and RPe-PL-D2 were identified up to the 24th by the sequencer. Four peptides (L1–L4) were isolated from the lysylendopeptidase digest of RPe-PL-D1 by RP-HPLC. By direct sequencing, peptides L1, L2, and L4 were sequenced completely, and peptide L3 up to the 30th. Peptide L3 was digested with trypsin or chymotrypsin and the resulting peptides were separated by RP-HPLC, yielding four (L3T1–L3T4) or five peptides (L3C1–L3C5). By sequencing them, the sequence of L3 was completed. RPe-PL-D1 was cleaved with CNBr in 70% formic acid by the method of Gross 99 and the resulting two fragments (CB1 and CB2) were isolated by RP-HPLC. From their N-terminal sequences, the linkages of CB1–CB2 and L3–L4 were connected. Thus, the complete amino acid sequence of PL-D1 was established. In the same way, the complete amino acid sequence of PL-D2 was also established. The sequence of PL-D2 was completely identical to that of PL-D1 except for the lack of C-terminal Leu-Thr.

Carboxypeptidase Y (Oriental Yeast Co., Ltd.) digestion of 1 nmol of reduced and S-carboxymethylated (RCm)-PL-D1, prepared by the method of Crestfield et al. 101 with about 1/5 (w/w) enzyme in 20 mM sodium phosphate buffer, pH 5.5, at 20 C released Thr (0.79 nmol) and Leu (0.45 nmol) at 1 min, Thr (0.95 nmol), Leu (0.95 nmol), Asp (0.07 nmol), and CmCys (0.04 nmol) at 30 min, and Thr (0.98 nmol), Leu (0.95 nmol), Asp (0.50 nmol), and CmCys (0.28 nmol) at 60 min, while that of 1 nmol of RCm-PL-D2 released Asp (0.10 nmol) at 20 min, Asp (0.23 nmol) and CmCys (0.09 nmol) at 30 min, and Asp (0.54 nmol) and CmCys (0.20 nmol) at 60 min. This result confirmed the C-terminal sequence of PL-D1 and PL-D2 to be -CDLT and -CD, respectively.

The high susceptibility of the C-terminal Leu-Thr of PL-D1 suggests the possibility that PL-D2 may be derived from PL-D1.

Amino acid sequence analysis of PL-D1 is summarized in Fig. 2. PL-D1 and PL-D2 were titrated with 5,5-dithio-bis(2-nitrobenzoic acid) in the presence of 6 M guanidine-HCl by the method of Ellman, 111 but no free sulfhydryl group was detected, suggesting that all cysteine residues were linked by disulfide bonds. PL-D1 and PL-D2 consist of 84 and 82 amino acid residues and their molecular masses were calculated to be 9317 and 9103 Da, respectively.

Comparison of the amino acid sequence of PL-D1 with those of PL-C,127 Urtica dioica agglutinin (UDA),128 and WGA2 129 is shown in Fig. 3. The amino acid sequence of residues 1-42 and residues 43-82 of PL-D1, designated domain A and domain B, have 50% homology with each other, and 80% homology to the domain A and domain C of PL-C, respectively. Twelve amino acid residues including eight cysteine residues (Cys4, Gly5, Cys13, Cys18, Cys19, Ser20, Gly23, Cys25, Gly26, Cys32, Cys36, and Cys40 in domain A) are absolutely conserved in these chitin-binding domains.

The hemagglutinating activity of PL-D was assayed after incubation of PL-D with 2% rabbit erythrocyte suspension in

---

Fig. 2. Summary of Amino Acid Sequence Analysis of PL-D1.

L1 and L2 indicate the peptides obtained by lysylendopeptidase digestion and CNBr cleavage of RPe-PL-D1, respectively. L3T and L3C indicate the peptides obtained by tryptic and chymotryptic digestions of peptide L3. Sequencing data under individual peptides are indicated as follows: , sequenced by the DABITC/PTC method; , sequenced by gas-phase sequencer; , sequenced by carboxypeptidase Y digestion; , not identified by sequencing.

---

Fig. 3. Comparison of Amino Acid Sequence of PL-D1 with Those of Pokeweed Lectin-C (PL-C),127 Urtica dioica Agglutinin (UDA),128 and Wheat Germ Agglutinin 2 (WGA2). 129

* stands for the domains of PL-D1, PL-C, UDA, and WGA2. Identical residues in the A and B domains of PL-D1 are bold-faced and those in all eleven domains are boxed. Several gaps (-) are inserted to obtain the maximum homology among these domains. <Q> pyroglutamate.
Fig. 4. UV-Difference Spectra of PL-D1 in the Presence of (GlcNAc)₄₋₄ (A) and Variation of the Different Molar Absorptions at 293 nm as a Function of the (GlcNAc)₄₋₄ Concentration (B).

(A): The difference spectra were measured at 20°C in 150 mM NaCl-50 mM phosphate buffer, pH 7.0. The concentrations of PL-D and (GlcNAc)₄₋₄ were 30 μM and 0.1 mM, respectively. (1), GlcNAc; (2), (GlcNAc)₂; (3), (GlcNAc)₃; (4), (GlcNAc)₄. 

(B): Different molar absorptions of PL-D1 at 293 nm were measured as above in the presence of various concentrations of (GlcNAc)₄₋₄ (△), GlcNAc (●), (GlcNAc)₂ (△), (GlcNAc)₃ (▲). (GlcNAc)₄.

0.15 M NaCl-10 mM phosphate buffer, pH 7.4, at 37°C for 1 h as described previously. Both PL-Ds did not agglutinate native erythrocytes, but toward trypsin-treated erythrocytes exhibited a weak agglutinating activity (15.6 μg/ml), which was about 0.10% and 1.56% those of WGA and PL-B, respectively. The hemagglutination of PL-D was inhibited by β(1→4) linked GlcNAc oligomers and the minimum concentrations required to inhibit the agglutination were 6.25 and 3.12 mM for (GlcNAc)₃ and (GlcNAc)₄, respectively, but GlcNAc and (GlcNAc)₂ did not inhibit the hemagglutination even at 25 mM.

To learn the interaction spectrum of PL-D with these saccharides, the ultraviolet (UV)-difference spectrum of PL-D1 (20 μM) induced by (GlcNAc)₄₋₄ in 0.15 M NaCl-50 mM phosphate buffer, pH 7.0, was examined by the method of Matsumoto et al. using a Shimadzu UV-2200 spectrophotometer. In the case of 0.1 mM (GlcNAc)₄₋₄, as shown in Fig. 4A, the UV-difference spectrum of PL-D was induced by (GlcNAc)₄ and (GlcNAc)₃ capable of inhibiting hemagglutination, but not by GlcNAc and (GlcNAc)₂, showing a non-inhibitory effect. The UV-difference spectra induced by (GlcNAc)₃ and (GlcNAc)₂ had a maximum at 292-294 nm and a smaller maximum at 284-285 nm, and their difference molar absorptions at 293 nm increased with increasing concentration of the saccharides to a plateau level (Fig. 4B). Such difference spectrum is ascribed to the red shift of tryptophan residue, suggesting that tryptophan residue(s) is at or near the saccharide-binding site of PL-D1.

References