Purification, Characterization, and Complete Amino Acid Sequence of a Trypsin Inhibitor from Amaranth (Amaranthus hypochondriacus) Seeds¹

Silvia Valdes-Rodriguez, Magdalena Segura-Nieto, Alicia Chagolla-Lopez, Aurora Verver y Vargas-Cortina, Norma Martinez-Gallardo, and Alejandro Blanco-Labra*

Department of Biotechnology and Biochemistry (S.V.-R., A.C.-L., N.M.-G., A.B.-L.) and Department of Genetic Engineering (M.S.-N., A.V.V.-C.), Centro de Investigacion y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Irapuato, Apdo. Postal 629, C.P. 36500, Irapuato, Guanajuato, Mexico

A protein proteinase inhibitor was purified from a seed extract of amaranth (Amaranthus hypochondriacus) by precipitation with (NH₄)₂SO₄, gel-filtration chromatography, ion-exchange chromatography, and reverse-phase high-performance liquid chromatography. It is a 69-amino acid protein with a high content of valine, arginine, and glutamic acid, but lacking in methionine. The inhibitor has a relative molecular weight of 7400 and an isoelectric point of 7.5. It is a serine proteinase inhibitor that recognizes chymotrypsin, trypsin, and trypsin-like proteinase activities extracted from larvae of the insect Prostephanus truncatus. This inhibitor belongs to the potato-I inhibitor family, showing the closest homology (59.5%) with the Lycopersicum peruvianum trypsin inhibitor, and (51%) with the proteinase inhibitor 5 extracted from the seeds of Cucurbita maxima. The position of the lysine-aspartic acid residues present in the active site of the amaranth inhibitor are found in almost the same relative position as in the inhibitor from C. maxima.

Amaranth is a very ancient crop. Its presence in Mexico dates from 4000 B.C. in Tehuacan, Puebla (Teutonico and Knorr, 1985), and thus it is one of the oldest known food crops. This grain was an important nutrient for the Aztec, Maya, and Inca civilizations. It was used to prepare a special bread used in religious ceremonies, and mainly for this reason, the crop was banned by the Spaniards when they arrived in America.

Three species of the genus Amaranthus produce edible seeds: Amaranthus hypochondriacus, grown in Mexico; Amaranthus cruentus, grown in Guatemala and Mexico; and Amaranthus caudatus, grown in Peru. Due to its high-quality protein, especially its relatively high Lys content (Downton, 1992), this crop has received considerable attention as a supplement to cereals and legumes to prevent protein malnutrition.

Enzyme inhibitors found in seeds have received particular attention because of their potentially deleterious effects in animal and human nutrition and their possible role in the defense of plants against microbial and insect pests. Some of these inhibitors have a double function against proteases and amylases (Richardson et al., 1987).

Trypsin inhibitors in amaranth seeds have been reported to range from 900 to 2100 units of activity for the black seeds, and from 300 to 5150 units for the yellow seeds (Duarte-Correa et al., 1986). Three functions have been proposed for plant proteinase inhibitors: as regulatory agents in controlling endogenous proteinases, as storage proteins, and as protective agents directed against insects or microbial proteinases (Ryan, 1973).

In this paper the specificity and the amino acid sequence of a trypsin inhibitor purified from amaranth seed extracts and its homology with some other inhibitors are reported.

MATERIALS AND METHODS

Seeds of Amaranthus hypochondriacus line 53 were kindly provided by the National Institute for Research in Forestry and Agriculture. Sephadex G-75 and DEAE Sepharose CL-6B were obtained from Pharmacia Fine Chemicals. All the reagents for electrophoresis were obtained from Bio-Rad. Bovine trypsin (EC 3.4.21.19) and chymotrypsin (EC 3.4.21.1), as well as BAEE, BAPNA, TAME, BTEE, and L-Leu p-nitroanilide, were supplied by Sigma. All enzymes used in the digestion of the inhibitor for sequencing were obtained from Boehringer Mannheim Biochemicals. All chemicals used were analytical grade.

Inhibitor Extraction

A. hypochondriacus seeds were ground into a fine powder by using a Wiley mill fitted with a 1-mm sieve. The powder (150 g) was defatted by stirring with 750 mL of a chloroform:methanol (2:1, v/v) mixture for 30 min. The extraction was repeated three times. The dry powder (100 g) was extracted at 4°C for 4 h with 500 mL of 0.02 M succinic acid buffer (pH 4.5), the suspension was clarified by centrifugation.

Abbreviations: ATI, amaranth trypsin inhibitor; BAEE, N-a-benzoyl-L-arginine ethyl ester; BAPNA, N-a-benzoyl-L-arginine-p-nitroanilide; BTEE, N-benzoyl-L-tyrosine-ethyl ester; GOR, Garnier Osguthorpe Robson; ITH5-CUCMA, trypsin and Hageman factor inhibitor from Cucurbita maxima; pl, isoelectric point; TAME, N-a-p-tosyl arginine methyl ester.

¹ This research was supported by the Organization of the American States (Multinational Project of Biotechnology and Food).

* Corresponding author; fax 52-462-51282.
at 10,000g for 1 h, and the pellet was discarded. The crude protein in the extract was precipitated with 35 to 65% saturation with ammonium sulfate. The precipitate was dissolved in water, dialyzed, and centrifuged at 17,000g for 20 min at 4°C and lyophilized. The lyophilized supernatant (300 mg) was dissolved in 5 mL of a 0.01 M ammonium bicarbonate solution and fractionated on a 2.5 × 165 cm Sephadex G-75 chromatography column, equilibrated with 0.01 M ammonium bicarbonate.

Fractions containing trypsin inhibitor activity were pooled and lyophilized. This material (250 mg) was purified on a CM-Sepharose CL-6B column, equilibrated with 0.02 M ammonium acetate (pH 6). The trypsin inhibitor was eluted with 500 mL of a linear gradient of ammonium acetate (0.02–1.0 M). After lyophilization, the active fractions were subjected to reverse-phase HPLC using a model 1050 Hewlett-Packard system and a preparative Vydac C18 column (ID 22.5 mm, length 250 mm with a 10-μm particle size). The two solvents used were 0.1% (v/v) TFA in water (solvent A) and 0.1% (v/v) TFA in acetonitrile (solvent B). During elution (flow rate 1.0 mL min⁻¹), a linear gradient in which the solvent composition changed from 0 to 70%B (v/v) in 75 min was used. The eluent was monitored at 216 nm and the active fractions were collected and lyophilized.

**Enzyme Inhibition Assay**

The inhibition of trypsin and trypsin-like activities were followed by monitoring the rate of hydrolysis of BAEE at 253 nm (Schwertz and Takenaka, 1955). Chymotrypsin inhibition was assayed using BTEE as substrate as described by Suk-Hee and Morton (1973). Inhibition of aminopeptidase activity was measured with L-Leu p-nitroanilide at 405 nm as described by Appel (1974). Inhibition of proteinases extracted from germinating amaranth, insect larvae (with the exception of that from Prosop stephanus truncatus), and fungi were measured by following the hydrolysis of casein (Kakade et al., 1969) or hemoglobin (Lenney, 1975). The substrate was added after the enzyme and inhibitor had been preincubated at 30°C for selected periods. One unit of proteolytic activity was defined as the increase of 0.1 absorbance units under the assay conditions. The inhibitory activity was obtained by measuring the difference between the proteolytic activity in the absence and presence of the inhibitor.

**Proteases from Germinating Amaranth**

Amaranth seeds were soaked in 1% (v/v) sodium hypochlorite solution for 10 min and rinsed with distilled water. The seeds were placed on a wet sponge covered with filter paper and incubated at 30°C. Germinating seeds were collected at 17, 20, 24, and 40 h and kept in liquid nitrogen. The seeds were homogenized with water and centrifuged. The supernatant was used as the source of amaranth proteases.

**Extraction of Larval Enzymes**

Defatted larvae of *Sitophilus zeamais*, *Tribolium castaneum*, *Callosobruchus maculatus*, and *Acanthoscelides obtectus* were extracted with 0.2 M succinic acid buffer (pH 4.5) in a 1:5 (w/v) ratio. In the case of *P. truncatus*, the extraction buffer was 0.04 M succinic acid and 0.1 M NaCl (pH 6.5). In all cases the suspensions were centrifuged at 10,000g for 10 min at 4°C. This supernatant served as the source of larval proteinases.

**Crude Fungal Enzymes**

The spores obtained from potato dextrose agar cultures of *Aspergillus niger* and *Aspergillus fumigatus* were suspended in 10 mL of sterile water and added as an inoculum to Erlenmeyer flasks containing corn meal medium prepared as described by Johnson and Curl (1972). After addition of 1000 spores per mL of medium, the suspension was incubated at 28°C for 48 h with continuous stirring (100 rpm).

Growing mycelia were separated by centrifugation at 10,000g, the enzymes were then extracted by homogenization with glass beads in 0.05 M Tris-HCl buffer (pH 8.6) in a 1:4 (v/v) ratio. The suspensions were centrifuged at 10,000g for 20 min, and the enzyme activities were obtained from the supernatant. The proteolytic activity was assayed at different pH values.

**Electrophoresis**

The purified trypsin inhibitor was subjected to electrophoresis using the method of Schagger and Von Jagow (1987). A gel with the following acrylamide concentration was used: 4% (w/v) stacking gel, 10% (w/v) spacer gel, and 16% (w/v) separating gel.

**pI Determination**

The pI of a 20-μg sample was measured in a nondenaturing gel using a pH gradient from 3.5 to 9.5. A 5% (w/v) polyacrylamide and 3% (w/v) cross-linker gel (110 × 85 × 1 mm) was run for 1 h at 1500 V and 25 mA. The electrode solutions were 1 M H3PO4 and 1 M NaOH.

**Protein Determination**

The protein concentration of all fractions eluted during the different chromatographic procedures was estimated by measuring the absorbance at 280 and 216 nm.

**Reduction and S-Carboxymethylation**

The native proteinase inhibitor was reduced and S-carboxymethylated in 6 M guanidine-HCl, 0.1 M Tris buffer (pH 8.6) as described by Crestfield et al. (1963).

**Amino Acid Analysis**

Proteins were hydrolyzed in 6 M HCl-1% phenol (v/v) at 110°C for 24 h. The amino acids in the hydrolysates were derivatized with phenyl-isothiocyanate and analyzed on a PicoTag C18 column (Bidlingmeyer et al., 1984).

**Primary Structure Determination**

The N-terminal sequence of 1 nmol of the carboxymethylated protein was first determined using repeated cycles of...
Edman degradation (Hunkapiller et al., 1983), with an Applied Biosystems gas-phase sequenator (model 470). The phenylthiohydantoin amino acids were identified in a C_{18} Nova Pack (Waters) column on a Waters chromatographer.

The reduced and carboxymethylated proteins were desalted by gel filtration on PD10 and purified by reverse-phase HPLC on a C_{18} column as mentioned above. The protein was then separately digested with the following enzymes: Staphylococcus aureus V8 protease (EC 3.4.21.19) in 0.1 M ammonium bicarbonate buffer (pH 8); chymotrypsin (EC 3.4.21.1) in 0.1 M Tris-HCl buffer (pH 8), containing 0.1 mM CaCl_{2}; and endoproteinase Lys-C in 25 mM Tris buffer, 1 mM EDTA (pH 8). In all cases the enzymes were used at 2% concentration with respect to ATI concentration and incubated at 37°C. The incubation time varied from 3.5 h (chymotrypsin), 6 h (V8 protease), to 17 h (endoproteinase Lys-C). The reactions were stopped with 0.1% (v/v) TFA, 6% (v/v) acetonitrile. The digested protein was either stored at −20°C or injected into a reverse-phase C_{18} column. The peptides were eluted with a linear gradient from 0 to 70% (v/v) acetonitrile in 0.1% (v/v) TFA during 75 min. The eluted peptides were collected and lyophilized, and approximately 0.1 to 1 nmol were used for each sequence analysis.

**Comparison of Amino Acid Sequences**

The amino acid sequence of ATI was compared with other proteins stored in the SwissProt data bank, in Geneworks, version 2.0.

**Prediction of Hydrophathy and Secondary Structure**

The hydrophathy profile was predicted by the method of Kyte and Doolittle (1982). The hydrophobicity values were calculated using a window of 11 residues (5 before and 5 after a given residue). The secondary structure prediction was obtained with the method of GOR (Garnier et al., 1978). The range used was eight before and eight after a given residue. Both algorithms are in the Geneworks program from Intelli-genetics, Inc. (version 2.0, Mountain View, CA).

**RESULTS AND DISCUSSION**

Most of the trypsin inhibitory activity from the crude seed extract was obtained by precipitation at 35 to 60% (w/v) (NH_{4})_{2}SO_{4} saturation. The precipitate was then passed through a Sephadex G-75 column where the inhibitory activity was eluted as a single peak. Further purification of the inhibitory activity was performed by CM-Sepharose CL-6B chromatography using a gradient of ammonium acetate. Most of the protease inhibitor was eluted with 0.5 M ammonium acetate. The active fractions were pooled, lyophilized, and purified by reverse-phase HPLC. A typical chromatogram is shown in Figure 1. The major peak eluting at 37% (v/v) acetonitrile had most of the trypsin inhibitory activity. Some of the smaller peaks also showed inhibitory activity; however, only the largest one was selected for further work.

The homogeneity of the purified inhibitor was shown by SDS-PAGE, where a single band with a mol wt of approximately 7000, was observed (data not shown). The purified inhibitor is focused as a single band, with a pl of 7.5. This result, as shown in Figure 2, could be indicative of the absence of isoforms. Nevertheless, as is mentioned below, there are two microheterogeneities in the primary sequence; however, in both cases the charges of the two changed residues were similar (Ser^{θ}/Tyr and Thr^{θ}/Tyr).

The ATI purified by HPLC was assayed against proteases from different sources. It was shown (Table I) that the inhibitor strongly inhibited the activity of bovine chymotrypsin and trypsin. Also, the protease activity from the insect P. truncatus, which has been shown to be a trypsin-like enzyme (Houseman and Thie, 1993), was inhibited. Neither of the crude fungal proteinases extracted from the mycelia of A. niger and A. fumigatus were inhibited. These fungal extract proteases, which are not typical trypsin-like enzymes, have an optimal activity at pH 8.6 and are unable to hydrolyze specific substrates for trypsin-like proteases such as BAEE, BAPNA, and TAME.

Several other proteases were not inhibited by the ATI, such as the enzymes extracted from the larvae of the other five insects tested, as well as the proteases extracted from the germinating amaranth seeds (Table I). In the latter case, the extracts showed proteolytic activity at pH 2.5, 6.5, and 7.6, as well as an aminopeptidase-type activity. However, none of these activities were affected by the addition of this inhibitor. These results suggest that ATI is not playing a role in the control of endogenous proteinases of the type that has been shown in other cultivars (Blanco-Labra and Iturbe-Chinas, 1981; Weselake et al., 1983). As for the insect-pro-
germinating amaranth fungi whose enzymes are known to differ from trypsin were against trypsin-like proteinases, since none of the insects or our hypothesis that ATI is a specific inhibitor directed mainly to man and Nielsen, 1987), and (Lemos et al., 1985; Silva and Xavier-Filho, 1991), (Wei-A. obtectus (Gatehouse maculatus and Asp and Cys proteinase from C peptidases have been reported from S. zeamais (Baker, 1982), was due to an Asp proteinase. In the other cases, amino acidases have been reported from S. zeamais (Baker, 1982), and Asp and Cys proteinase from C. maculatus (Gatehouse et al., 1985; Silva and Xavier-Filho, 1991), A. obtectus (Weiman and Nielsen, 1987), and Zabrotes subfasciatus (Lemos et al., 1990; Silva and Xavier-Filho, 1991). These data confirm our hypothesis that ATI is a specific inhibitor directed mainly against trypsin-like proteinases, since none of the insects or fungi whose enzymes are known to differ from trypsin were inhibited by ATI. Therefore, it becomes evident that to think of using this type of inhibitor as a possible tool to improve plant resistance to insect or microbial attack, it is necessary to consider in advance the specific types of enzymes of the particular target organism (Christeller et al., 1992).

Analysis of the peptides derived from ATI by digestion with chymotrypsin, S. aureus V8 protease, and a Lys-specific proteinase enabled the deduction of the sequence shown in Figure 3. Redigestion of the endoproteinase Lys-C peptide 6 EC with the S. aureus V8 proteinase was helpful in confirming the sequence in the C-terminal region. Digestion with carboxypeptidase Y (EC 3.4.16.1) provided the sequence of the last two residues, i.e. -Val-Thr. The inhibitor contains 69 amino acids, corresponding to a M„ of 7400, which is in agreement with the mol wt estimated by SDS-PAGE. The amino acid sequence was also compatible with the amino acid composition (data not shown) showing a high Val, Arg, and Glu content.

By comparing the sequence of ATI to those of other proteins in the SwissProt data bank, it was shown that six polypeptides, all of which are protease inhibitors belonging to the potato I family, had the closest homology to ATI (Fig. 4). The two polypeptides with the greatest similarity to ATI were the Lycopersicum peruvianum inhibitor (59.5%), and the Cucurbita maxima trypsin inhibitor (51.0%). These two contained the same residues Lys-Asp, which have been reported to be an important part of the reactive site in the ITH5-CUCMA (Krishnamoorthi et al., 1990).

When comparing the structure of the ATI with that of the ITH5-CUCMA and with that of L. peruvianum, not only were the Lys-Asp fragments from the active site similar, but also the nine amino acids in which the Lys-Asp sequence occurs (44–53 for ATI) were almost identical. This nine-amino acid sequence could perhaps constitute the exposed binding loop of this type of inhibitor. Bode and Huber (1992) reported that most of the Ser protease inhibitors react with cognate enzyme, according to a common, substrate-like standard mech-

![Figure 2. Polycrylamide gel IEF of the ATI. Lane 1, Mixture of pl standard proteins; phycocyanin (pl 4.65), β-lactoglobulin B (pl 5.1), bovine carbonic anhydrase (pl 6.0), human carbonic anhydrase (pl 6.5), equine myoglobin (pl 7.0), human hemoglobin A (pl 7.1), human hemoglobin C (pl 7.5), lentil lectin (pl 7.8, 8.0, 8.2), Cyt c (pl 9.6). Lane 2, ATI purified by reverse-phase HPLC.](image)

**Table I. Effect of ATI against proteases from different sources**

The HPLC-purified inhibitor was assayed against each enzyme by at least duplicate determinations. Each value is the mean of three replicates of one of these experiments. +, Inhibition; –, no inhibition.

<table>
<thead>
<tr>
<th>Protease</th>
<th>Inhibition</th>
<th>UI µg⁻¹ ATI ± SD*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bovine chymotrypsin</td>
<td>+</td>
<td>2.65 ± 0.12</td>
</tr>
<tr>
<td>Bovine trypsin</td>
<td>+</td>
<td>2.59 ± 0.08</td>
</tr>
<tr>
<td>Germinating amaranth</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Insects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prostephanus truncatus</td>
<td>+</td>
<td>2.80 ± 0.05</td>
</tr>
<tr>
<td>Sitophilus zeamais</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Tribolium castaneum</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Callosobruchus maculatus</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Zabrotes subfasciatus</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Acanthoscelides obtectus</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Fungi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspergillus niger</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Aspergillus fumigatus</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

* Units of inhibition per µg of ATI.

![Figure 3. Amino acid sequence of ATI. The arrows indicate the polypeptides obtained and sequenced after each enzymic digestion. Ch, Chymotryptic peptides; V, peptides from the digestion with protease from S. aureus V8; EC, peptides from digestion with endoproteinase Lys-C. Amino acid residues were identified by automated sequencing.](image)
anism. They all present an exposed binding loop of a characteristic canonical conformation, but are otherwise unrelated in structure.

The three-dimensional structures of two inhibitors belonging to the potato I inhibitor family have been reported: barley chymotrypsin inhibitor-2 (McPhalen et al., 1985), and eglin-c inhibitor. In the case of eglin-c inhibitor, the structure of the x-ray structure at 1.95 Å of the free eglin-c inhibitor. Both inhibitors lack disulfide bridges.

In the case of ATI there are two Cys’s in very similar positions to those of the ITH5-CUCMA, 3 and 48 and 4 and 49 for ATI. However, it is not known whether those Cys’s are involved in a disulfide linkage.

The hydrophatic character of ATI was analyzed through the computer program of Kyte and Doolittle (1982) by progressively evaluating the hydrophobicity of this protein along its amino acid sequence. In Figure 5, the relatively more hydrophobic region that occurs at the C terminus, starting from residue 52, is shown. In analogy to the basic pathogenesis-related proteins (Huib and Linthorst 1991), the hydrophobic domain at the C-terminal sequence could be involved in intracellular targeting. Finally, the secondary structure prediction using the GOR algorithm (Garnier et al., 1978) suggests that in the first two-thirds of the protein there are four regions with a propensity to form β sheets. This information should contribute to knowledge about the structure and the interaction of these types of protease inhibitors with insect and microbial enzymes. By learning its sequence and comparing it with some other sequences whose three-dimensional structures have already been determined, it is possible to suggest a similar structure for the ATI. The characterization of the amaranth inhibitor’s activity against trypsin and trypsin-like proteinases and chymotrypsin. This property could be useful for crop improvement against insect attack and could contribute toward better understanding of the possible role of some of the components involved in the defense mechanism of plants.

ACKNOWLEDGMENTS

We are grateful to INIFAP (National Institute for Forestry and Agriculture, Celaya Unit) for supplying the amaranth seeds. We also thank Yolanda Rodriguez, Elizabeth Mendiola, Armando Guerrero, and Ana L. Ruiz for their valuable technical assistance, June Simpson for her critical review of the manuscript, and R. Rivera for his assistance in the use of the computer programs. Received May 17, 1993; accepted August 9, 1993.

LITERATURE CITED


Baker JE (1982) Digestive proteases of Sitophilus weevil (Coleop-


Schägger H, Von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal Biochem 166: 368–379


Silva CP, Xavier-Filho J (1991) Comparison between the levels of aspartic and cysteine proteinases of the larval midguts of Callosobruchus maculatus (F.) and Zabrottes subfasciatus (Boh.) (Coleoptera:Bruchidae). Comp Biochem Physiol 99B: 529–533


