A Novel α-Amylase Inhibitor from Amaranth (Amaranthus hypocondriacus) Seeds*

(Received for publication, May 4, 1994, and in revised form, June 14, 1994)

Alicia Chagolla-Lopez‡§*, Alejandro Blanco-Labran†, András Pathy§, Roberto Sánchez++, and Sándor Pongor°‡¶

From the $International Centre for Genetic Engineering and Biotechnology, 34012 Trieste, Italy, §Cinvestav IPN ap 629 Irapuato Gto. Mexico, and Institute of Biochemistry and Protein Research, Agricultural Biotechnology Center, 2101 Godollo, Hungary

The major α-amylase inhibitor (AAI) present in the seeds of Amaranthus hypocondriacus, a variety of the Mexican crop plant amaranth, is a 32-residue-long polypeptide with three disulfide bridges. Purified AAI strongly inhibits the α-amylase activity of insect larvae (Tribolium castaneum and Prostephanus truncatus) and does not inhibit proteases and mammalian α-amylases. AAI was sequenced with the automated Edman method, and the disulfide bridges were localized using enzymatic and chemical fragmentation methods combined with N-terminal sequencing. AAI is the shortest α-amylase inhibitor described so far which has no known close homologs in the sequence data bases. Its residue conservation patterns and disulfide connectivity are related to the squash family of proteinase inhibitors, to the cellulose binding domain of cellobiohydrolase, and to o-conotoxin, i.e., a group of small proteins termed "knotting" by Nguyen, D. L., Heitz, A., Chiche, L., Castro, B., Boige-grain, R., Favel, A., and Coletti-Previero, M. (11990) (Biochimie 72, 431–435) The three-dimensional model of AAI was built according to the common structural features of this group of proteins using side-chain replacement and molecular dynamics refinement techniques.

Enzyme inhibitors are important tools of nature for regulating the activity of enzymes in cases of emergency. Plant seeds are known to produce a variety of enzyme inhibitors that are thought to protect the seed against insects and microbial pathogens. Proteinase inhibitors are the best studied of this group (1); expression of proteinase inhibitor genes in transgenic plants provides protection against pathogens (for a review, see Ryan (2)). Comparatively less is known about the inhibitors of α-amylase which might, on the other hand, be equally attractive can-indicate this fact. For example, members of the cereal family of amylase/protease inhibitors are active against insect α-amylases but do not seem to inhibit the α-amylases present in the digestive system of mammals.

Here we report on the purification of a new type of α-amylase inhibitor isolated from the seeds of Amaranthus hypocondriacus which strongly inhibits the α-amylase of the larvae of the red flour beetle (Tribolium castaneum) and of the grain borer (Prostephanus truncatus). The primary structure of this small protein is not closely related to any other known protein. Its disulfide topology and residue conservation patterns, however, are similar to those of a group of proteins that include members of the squash family of proteinase inhibitors (1, 18, 19), as well as o-conotoxins (20, 21), which contain three conserved disulfide bridges and an array of three β sheets.

EXPERIMENTAL PROCEDURES

Materials—Seeds of A. hypocondriacus line 55 were kindly provided by the Mexican National Institute for Research in Forestry and Agriculture in Celaya Gto., Mexico. Sephadex G-75 and DEAE-Sepharose CL-6B were obtained from Pharmacia Biotech Inc. α-Chymotrypsin and trypsin were from Serva; cyanogen bromide and vinyl pyridine were from Aldrich. All chemicals used were of analytical or sequencing grade. HPLC-grade acetonitrile and trifluoroacetic acid were obtained from Aldrich.

α-Amylase Assay—Crude extracts of larval α-amylases of T. casta-neum and P. truncatus were extracted as described previously (22). The activity of α-amylase was determined using nitrosalicylic acid according to the Bernfeld method (8).

Purification of the α-Amylase Inhibitor—A crude extract of α-amylase inhibitor was obtained from 100 g of ground defatted amaranth seeds as described previously (22). The supernatant was precipitated by the addition of ammonium sulfate. The precipitate in the range of 35–65% saturation was collected and redissolved in 0.01 M ammonium bicarbonate. This material was fractionated on a Sephadex G-75 column (1.6 × 160 cm) precolumn and with the same solution. Fractions inhibiting the α-amylase activity of T. castaneum larvae were pooled and lyophi-lized. This material was redissolved in 5 ml of 0.02 M ammonium bicarbonate buffer, pH 8.3, and chromatographed on a (20 × 2.6 cm) DEAE-Sepharose CL-6B column precolumn and with the same buffer using a linear gradient of ammonium bicarbonate (0.02-0.5 M). The active fractions eluted were lyophilized and subjected to reverse phase HPLC on a preparative Vidac C18 column (22.5 × 250 mm, 10 mm particle size) using a model 1050 Hewlett Packard HPLC system. The two solvents used were 0.1% (v/v) trifluoroacetic acid in water (solvent A) and 0.1% (v/v) trifluoroacetic acid in acetonitrile (solvent B). A linear gradient was used for elution (flow rate. 5 ml/min) in which the solvent composition changed from 0 to 80% B in 120 min.

Amino Acid Analysis—Amino acid composition of the peptides was determined using a Waters workstation and PicoTag HPLC system

* 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.

§ Recipient of an International Centre for Genetic Engineering and Biotechnology fellowship.

** Recipient of an International Centre for Genetic Engineering and Biotechnology fellowship.

† To whom correspondence should be addressed. Tel.: 39-40-3757300; Fax: 39-49-226-555; E-mail: pongor@icgeb.trieste.it.
A Novel Amaranth Seed α-amylase Inhibitor

TABLE I

Table: A classification of α-amylase inhibitors

<table>
<thead>
<tr>
<th>Class</th>
<th>Source*</th>
<th>Size [aa]</th>
<th>1/2-Cys</th>
<th>Against in-</th>
<th>Against mam-</th>
<th>Against pro-</th>
<th>Members of the group with other activities†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kunitz type</td>
<td>Barley (4), wheat (5), rice (6)</td>
<td>176–180</td>
<td>2–4</td>
<td>+ + +</td>
<td>+</td>
<td>+</td>
<td>Miracin (7)</td>
</tr>
<tr>
<td>Cereal type</td>
<td>Wheat (3), barley (3), Indian finger millet (8)</td>
<td>124–160</td>
<td>10</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>ND*</td>
</tr>
<tr>
<td>γ-Purothionin type</td>
<td>Sorghum (9)</td>
<td>47–48</td>
<td>8</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>γ-Purothionin (10)</td>
</tr>
<tr>
<td>Ragi I-2 type</td>
<td>Indian finger millet (11)</td>
<td>95</td>
<td>7</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>Phospholipid transfer proteins</td>
</tr>
<tr>
<td>Legume lectin type</td>
<td>Common beans (12)</td>
<td>246</td>
<td>10–16</td>
<td>+ +</td>
<td>+</td>
<td>+</td>
<td>Legume lectins (13)</td>
</tr>
<tr>
<td>Thaumatin type</td>
<td>Maize (14)</td>
<td>173–235</td>
<td>4</td>
<td>+ +</td>
<td>–</td>
<td>–</td>
<td>Pathogenesis related protein (15) osmotin (16) thauatin (17)</td>
</tr>
<tr>
<td>Prokaryotic</td>
<td>Actinomycetes</td>
<td>75–120</td>
<td>4</td>
<td>+ +</td>
<td>–</td>
<td>–</td>
<td>ND</td>
</tr>
</tbody>
</table>

* Numbers in parentheses are references.
† ND, no data.

**Table**: A classification of α-amylase inhibitors

A nonbonded cutoff value of 10 Å was used together with a distance-dependent dielectric constant. Dynamics was performed using a time step of 1 fs, and the temperature was kept constant by coupling to a thermal bath with a constant of 0.1 ps (34). The refinement was a modification of the procedure described by Du and co-workers (35). An initial minimization was performed with 200 cycles of SD minimization followed by 200 cycles of CG minimization. Afterward, the model was put into a 50-Å diameter sphere of pre-equilibrated water and the system was subjected to 5 ps dynamics at 300 K and 200 cycles of SD minimization keeping the protein atoms fixed, in order to allow badly placed water molecules to move. This was followed by a 5 ps dynamics at 200 K keeping only the SCRs fixed. The refinement proceeded with a low temperature simulated annealing starting at 340 K and lowering the temperature to 260 K using 20-K steps. The system was kept for 5 ps at each temperature. The system was minimized using 200 cycles of SD minimization and 200 cycles of CG minimization. Finally, the protein-solvent system was subjected to a 100-ps molecular dynamics simulation at 300 K, allowing all atoms to move. The resulting structure was subjected to 200 steps of SD energy minimization and to a CG minimization until the maximum energy derivative was lower than 0.1 kcal/A.

The model was evaluated using the Prosai program (36). The program uses mean force potentials derived from known protein structures. Two kinds of potentials, Cβ-Cβ pair interaction potentials and surface potentials, are used for this purpose. As the surface potentials are not recommended for small proteins (37) (in fact the surface potential calculations failed to recognize the native structure in 8 out of 10 small disulfide-rich proteins tested by us), we based the analysis of the AAI structure solely on the Cβ-Cβ interaction potentials. The program was used with standard parameters as provided by the author.

**RESULTS**

Purification of A. hypocondriacus α-Amylase Inhibitor—The crude extract prepared by succinic acid extraction contained inhibitors of both α-amylase and trypsin (data not shown). When subjected to gel filtration on a Sephadex G-75 column, the fractions inhibiting α-amylase of the larval enzyme of T. castaneum eluted essentially as a single peak. These fractions were lyophilized and subjected to anion exchange chromatography on a DEAEP-Sepharose CL-6B column. No inhibitory activity was found in the unbound fraction and a linear NH₄HCO₃ gradient allowed for the separation of two peaks showing inhibitory activity (IEX-1 (left) and IEX-2 (right), in Fig. 1). The respective fractions were pooled and subsequently subjected to reverse phase HPLC. IEX-1 yielded several peaks with α-amylase inhibitory activity that were not analyzed further in this study. Reverse phase HPLC of the IEX-2, on the other hand, gave one major peak eluting at 25% acetonitrile, which we called amaranth α-amylase inhibitor (AAI), and subjected to sequencing.
Specificity of AAI—Crude extracts of amaranth seeds inhibit both trypsin (22) and insect amylases. Purified AAI showed inhibitory activity against the larval α-amylase of T. castaneum and that of P. truncatus. On the other hand, AAI does not inhibit human or bovine saliva α-amylases in an appreciable manner (the species-specificity studies will be published elsewhere).

Amino Acid Sequence and Disulfide Topology of AAI—As the protein showed a high percentage of cysteine with no free sulfhydryl groups, the samples were subjected to reduction and pyridylethylation. Digestion of the reduced and pyridylethylated protein with trypsin and cyanogen bromide resulted in seven overlapping peptides that were separated by reverse phase HPLC and sequenced with automatic Edman degradation. Notable features of the AAI sequence (Fig. 2) are the high content of cysteine (six) and proline (four) within 32 residues.

Disulfide bridges were determined through partial double digestion with trypsin/chymotrypsin at pH 2.0 for 4 h. On reverse phase HPLC, the digestion mixture yielded six main peaks (TCA1–TCA6), one of them (TCA5) being identical to the native protein according to amino acid composition. The other peaks were analyzed through sequencing as well as by amino acid analysis, the deduced sequences are shown in Fig. 3A. The results indicate that Cys⁸ and Cys²⁹ are connected by a disulfide bond. Furthermore, both Cys⁴ and Cys¹⁷ form a disulfide bond either with Cys¹⁷ or with Cys²⁹, respectively. The exact placement of these disulfide bonds could not be determined from these experiments since the enzymatic cleavage of the Cys¹⁷-Cys²⁹ peptide bond cannot be accomplished. The two theoretically possible disulfide bonding patterns are shown in Fig. 3B.

Sequence and Structural Similarities—When the sequence was compared with all known proteins using the FASTA (38), the BLAST (28) and the BLITZ (29, 30) programs, no convincingly significant homologies could be detected and the top list of homologies substantially varied depending on the choice of search parameters (data not shown). It has to be mentioned...
A Novel Amaranth Seed α-Amylase Inhibitor

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WNR</td>
</tr>
<tr>
<td>2b</td>
<td>CGPK TCTSDY</td>
</tr>
<tr>
<td>3ac</td>
<td>CIPK MDGVPCCEPY GHCK</td>
</tr>
</tbody>
</table>

![Fig. 3. Determination of the disulfide bonds in AAI by partial enzymatic digestion.](image)

that behavior is characteristic of short and compositionally biased query sequences such as AAI (39). In order to increase the sensitivity of the search, we selected a subset of the data base in which cysteine residues were distributed in a way similar to AAI. This search allowed us to tentatively identify a group of short cysteine-rich proteins and protein domains of different organisms, including various carbohydrate-binding proteins (celllobiohydrolase, wheat germ agglutinin, hevein, chitinase), toxins (conotoxins), antimicrobial peptides from *Amaranthus* (AMP), and the sweet taste-suppressing protein, guframin (Fig. 4). The best similarity (46% residue identity) was found in the case of the cellulose binding domain of celllobiohydrolase II from *Trichoderma reesei*. The homologous cellulose binding domain of celllobiohydrolase I is also included in the alignment. In this structure one disulfide bridge (denoted a in Fig. 3) is missing.

Given that the three-dimensional structure of several of these proteins is already known from x-ray and NMR studies, we could classify the structures in two groups based on disulfide topologies (Fig. 5, top) and folding patterns (Fig. 5, bottom). Group I contains the squashed family of trypsin inhibitors, the cellulose binding domain of celllobiohydrolase and o-conotoxin (1cc, 1chb, 1cti, 2eti). As the structures in this family contain a characteristic knot-like arrangement, Nguyen et al. (47) suggested the “knotting” name for this family of proteins. Group II contains chitin-binding domains of wheat germ agglutinin and hevein (1hev and 9gwa). In group I (topologyabcabc in Fig. 5, top) there are three disulfide bridges, whereas in group II (topologyabcacbd in Fig. 5, top), there are four. When viewed in the sequence context, three of these are in a topological arrangement seen in group I (i.e. abcabc), and one bridge d is outside. A comparison of the available three-dimensional structures revealed that a sheet composed of three short β-strands is present in both structural groups (Fig. 5, bottom). The disulfide bridges are, however, arranged in a different manner within the two groups. In group I all three disulfide bridges take part in the “reinforcement” of the sheet structure. In group II the three short β-strands are present but bridge c is connected to a short helical segment (absent in group I) connecting two strands of the sheet, while the fourth bridge d connects two ends of the C-terminal strand (symmetry shown in Fig. 5, top).

The common element of the two folding patterns is a short β-hairpin-like structure with an irregular N-terminal extension (boxed in Fig. 5, bottom). The three-strand arrangement is seemingly common to both folding patterns, however, while the third strand is located at the N terminus of the common pattern in group I, in the group II structures it is at the C terminus. Though the strands of the sheet are short, the three-strand arrangement can be seen on all but one of the structures related to AAI. The only exception is the carboxypeptidase A inhibitor (4cpa), in which no regular secondary structures can be detected even though its overall folding pattern is clearly related to the other structures in group I.

**Modeling**—The superimposition of the structures and the alignment of the structurally conserved regions was used to design a structural template for AAI. Assignment of two disulfide bridges in AAI was, in principle, an open question, since the connectivity of adjacent cysteines 17 and 18 could not be directly determined by chemical means. Theoretically there are two possibilities to form SS bridges using our connectivity data (Fig. 3B). (i) The “abcabc” topology is characteristic of group I structures, and was also found by chemical means in o-conotoxin (48); (ii) The abacbc topology, on the other hand, has not yet been found experimentally in short proteins (49). On the basis of the chemical evidence (48) and of the convincing similarity of AAI to group I proteins (Fig. 4), we chose the abacbc topology for our modeling studies.

We made a structural template in which the conformation of the first amino acids of (CPFILMR) followed that of the ET1 structure, while the rest of the molecule was modeled on the cellulose binding domain of T. reesei. Since the three-dimensional coordinates of the celllobiohydrolase II cellulose binding domain are not published, we used the structure of the celllobiohydrolase I cellulose binding domain (1chb (40)), which is reported to be identical with the former. To build this structure we had to introduce deletions in the (Thr17-Val18) and (Val25-Leu26) positions, respectively (numbering of the 1chb structure). The sequence alignment between the AAI and the template (Fig. 6) resulted in sequence identity for 11 out of 32 residues (34%). The model was finally constructed through a residue by residue replacement. The disulfide bridge originally absent in the 1chb framework was built and the model was refined by energy minimization and molecular dynamics to give the structure shown in Fig. 7.

The reliability of the model was tested by the knowledge-based template field approach of Sippl, as implemented in the Prosali (36, 61) program. The program calculates the Cα-Cα pair interaction energy for each residue in the sequence, and correctly folded proteins produce smooth energy plots with negative values (36). The AAI model gave an energy profile with values corresponding to those of native structures (Fig. 8). The energy profile had no positive regions that would indicate misfolded parts in the model. Also, the so-called Z-score or normalized energy value (36) was −3.9, which is within the range of the values expected for native proteins of this length (36). When tested with the AAI sequence, all other known structures gave higher Z scores indicating that the model presented here fits the sequence better than any of the other structures (not shown).

**DISCUSSION**

*Amaranthus* seeds seem to contain a number of α-amylase inhibitors that can be separated by ion-exchange chromatography and reverse phase HPLC. Here we report on the purification and the structure of AAI, the most abundant α-amylase inhibitor of amaranth seeds that accounts for more than half of inhibitory activity measurable in crude extracts. This protein shows strong α-amylase inhibitory activity against one of the most important pests of maize, *P. truncatus* (larger grain borer), and a pest of wheat flour, *T. castaneum* (rust-red beetle).
A Novel Amaranth Seed α-Amylase Inhibitor

Fig. 4. Sequence alignment of AAI with protein sequences of similar cysteine patterns. The sequences are taken from the Swiss-Prot data bank (44) as follows: Celllobiohydrolase II, (exoglucanase II, EC 3.2.1.91) from T. reesei, GUX2, TRIRE, 27-63; Cellobiohydrolase I, exoglucanase I, (EC 3.2.1.91) from T. reesei (40), GUX1, TRIRE, 479-593; ω-conotoxin GVI A, snake venom of Conus geographus, CXOG, CONGE, 46-73; ω-conotoxin MVIIA, snake venom of Conus magus (41), CXOA, CONMA, 1-25; Gurmarin, sweet taste-modifying protein from Gymnema sylvestre (43), AMF, MIRJA, 1-37; AMP, antimicrobial peptide from Amaranthus caudatus, AMP, MIRJA, 1-30; AMP1, antimicrobial peptide from Amaranthus caudatus, AMP, MIRJA, 1-36; AMP2, antimicrobial peptide from Amaranthus caudatus, AMP2, MIRJA, 1-36; Hevein, chitin-binding lectin from Hevea brasiliensis, HEVE1, HEVBR, 18-60; Chitinase, basic endochitinase I from rice (45), CHI1, ORYS, 19-61; WGA, wheat germ agglutinin (46), AGH1, WHEAT, 27-69 (domain 1), 70-112 (domain 2).

Fig. 5. Disulfide patterns (top) and folding pattern (bottom) of proteins distantly related to AAI I and II correspond to the groups shown in Fig. 4. The disulfide and folding patterns were deduced from known structural data in each group.

Fig. 6. The structural template used for the modeling AAI. eti is the Ecballium elaterium trypsin inhibitor structure (50), cbh is the cellulose binding domain of cellobiohydrolase I (40).

The inhibitory activity, according to our preliminary results, is absent or minimal against human or bovine α-amylases, thus this protein appears to be an ideal candidate for conferring insect resistance upon transgenic plants.

AAI is a 32-residue peptide containing 6 cysteines. The first residue of the sequence is not methionine, therefore AAI is probably synthesized as part of a larger precursor. The sequence of AAI shows no obvious similarities with any of the known proteins. Spurious similarities and an examination of the residue conservation pattern allowed us to identify a group of structurally related proteins which contains sugar binding proteins (wheat germ agglutinin and celllobiohydrolase), venom (ω-conotoxin), and antimicrobial peptides. Using the known three-dimensional structures we built a model based on the similarity of AAI to the squash family of trypsin inhibitors, ω-conotoxin, and to the cellulose binding domain of celllobiohydrolase. Nguyen and associates introduced the term "knottin" for this group of structures, based on a "knotlike" feature in the three-dimensional fold (47). This feature is retained in the model of AAI. AAI is the first α-amylase inhibitor described in this group.
A novel Amaranth Seed α-Amylase Inhibitor

Fig. 8. Energy plot for the refined AAI model as generated with the ProsaII (36) program. The curves are smoothed (thin line, window size of 5; thick line, window size of 13). Energies are represented in units of E/kT.

AAI seems to be the shortest of the peptide α-amylase inhibitors described so far and, in spite of its overall similarity to the squash family of proteinase inhibitors, AAI does not seem to inhibit proteases. Even though AAI has some potential similarity to other small proteins, which allowed us to build a three-dimensional model of this inhibitor, we tend to believe that the similarities are structural rather than evolutionary. In other words, short peptides may not have too many stable conformations for accommodating three disulfide bridges, therefore a similar fold may arise as a result of convergent evolution. Finally we mention that AAI seems to be a good core structure for protein engineering studies since several of the related proteins are known to be stable and to refold correctly from the reduced state in vitro (47).

Acknowledgments—We thank Dr. Manfred Sippel for providing us with the PROSA II program. The protein sequencing laboratory is supported by the endowment grant of the Agricultural Biotechnology Center, Gödöllő, Hungary.

REFERENCES