α-Amylase inhibitor (AAI), a 32-residue miniprotein from the Mexican crop plant amaranth (Amaranthus hypochondriacus), is the smallest known α-amylase inhibitor and is specific for insect α-amylases (Chagolla-Lopez, A., Blanco-Labra, A., Patthy, A., Sanchez, R., and Pongor, S. (1994) J. Biol. Chem. 269, 23675–23680). Its disulfide topology was confirmed by Edman degradation, and its three-dimensional solution structure was determined by two-dimensional 1H NMR spectroscopy at 500 MHz. Structural constraints (consisting of 348 nuclear Overhauser effect interproton distances, 8 backbone dihedral constraints, and 9 disulfide distance constraints) were used as an input to the X-PLOR program for simulated annealing and energy minimization calculations. The final set of 10 structures had a mean pairwise root mean square deviation of 0.32 Å for the backbone atoms and 1.04 Å for all heavy atoms. The structure of AAI consists of a short triple-stranded β-sheet stabilized by three disulfide bonds, forming a typical knottin or inhibitor cystine knot fold found in miniproteins, which binds various macromolecular ligands. When the first intercystine segment of AAI (sequence IPKWNR) was inserted into a homologous position of the spider toxin Huwentoxin I, the resulting chimera showed a significant inhibitory activity, suggesting that this segment takes part in enzyme binding.

Plant seeds produce a large variety of enzyme inhibitors that are thought to provide protection against insects and microbial pathogens. As plant seed inhibitors are often species specific, i.e. they inhibit enzymes of a well defined group of pathogenic organisms but do not affect the mammalian counterpart, they make attractive candidates for conferring pest resistance to transgenic plants (for a review see Ref. 1).

The α-amylase inhibitors vary considerably in their structures. Many of their structural relatives, e.g. proteinase inhibitors, osmotin, and salt-induced proteins (Table I), play roles in plant stress response. The smallest of the known α-amylase inhibitors, AAI, 1 is found in the seeds of Amaranthus hypochondriacus, a variety of the Mexican crop plant amaranth or Prince’s feather (2). AAI is a 32-residue polypeptide with three disulfide bridges, which has no significant sequence similarity to other proteins in the data bases. It has a spurious sequence similarity to various members of the so-called knottin (3) or “inhibitor-type cystine knot” (4) family, which includes various proteinase inhibitors and toxins. AAI is species specific; it inhibits α-amylase of several pathogenic insect larvae (Triblium castaneum, Prosthanus truncatus, Periplaneta americana, and Tenebrio molitor) but not the mammalian α-amylases.

Here we report the three-dimensional structure of AAI as determined by NMR spectroscopy and show via amino acid replacement and chimera construction that a short segment of the first loop of AAI is involved in enzyme inhibition. 2

EXPERIMENTAL PROCEDURES

Materials

AAI was prepared as described (5). Sephadex G-75 and DEAE-Sepharose CL6B were obtained from Amersham Pharmacia Biotech. α-Chymotrypsin, endoproteinase Glu-C, and trypsin were obtained from Sigma; cyanogen bromide and vinyl pyridine were obtained from Aldrich. All chemicals used were of analytical or sequencing grade. HPLC grade acetonitrile and trifluoroacetic acid were obtained from Aldrich.

Assay of α-Amylase Inhibition

Crude α-amylase from T. molitor and P. americana larvae was extracted and partially purified as described (6). Assays of α-amylase inhibition were performed according to Bernfeld (7) as described (5).

Peptide Mapping and N-terminal Sequencing

A 75-μg sample of AAI was dissolved in 150 μl of 0.2 M Tris-HCl buffer, pH 7.3, and digested with a mixture of trypsin (5 μg), chymotrypsin (3 μg), and endoproteinase Glu-C (3 μg) at 37 °C for 15 h. The mixture was separated by reverse phase-HPLC on a Vydac C18 column (2.1 × 250 mm). Selected peaks were collected, lyophilized, and covalently coupled to amophenyl glass beads in prepacked capillary
Solution Structure of the \(\alpha\)-Amylase Inhibitor AAI

**Table I**  
Structural classification of \(\alpha\)-amylase inhibitors  
Based on a classification by Richardson (33) and completed with recent data (2). ND, no data.

| Class          | Source                                      | Size (amino acids) | \(\%\) Cys | Inhibitory activity | Members of the group with other activities |
|----------------|---------------------------------------------|--------------------|------------|---------------------|--------------------------------------------|
|                |                                             |                    |            | Against insect      | Against mammalian or other amylases        | Against proteases                           |
|                |                                             |                    |            | amylases            |                                            |                                            |
|                |                                             |                    |            | +                   | +                                          | ND                                         |
| Kunitz type    | Barley (29), wheat (30), rice (6)          | 176–180            | 2–4        | +                   | +                                          | + Miraculin (32)                           |
| Cereal type    | Wheat (33), barley (33), Indian finger millet | 124–160            | 10         | +                   | –                                          | ND                                         |
| \(\gamma\)-Purothionin type | Sorghum (35)                                  | 47–48              | 8          | +                   | –                                          | – \(\gamma\)-Purothionins (36)             |
| Rag1–2 type    | Indian finger millet (37)                   | 95                 | 7          | +                   | –                                          | Phospholipid transfer proteins (39)        |
| Legume lectin type | Common beans (38)                          | 246                |            | +                   | –                                          | Legume lectins (39)                        |
| Thaumatin type | Maize (40)                                   | 173–235            | 10–16      | +                   | +                                          | Pathogenesis-related protein (41), Osmotin (42), Thaumatin (43) |
| Knottin type   | Amaranth (this work)                        | 32                 | 6          | +                   | –                                          | Proteinase inhibitors, neurotoxins (2)     |
| Prokaryotic    | Actinomyces                                 | 75–120             | 4          | +                   | +                                          | ND                                         |

**Table II**  
Amino acid sequences and inhibitor properties of the peptides used in this study

| Peptide | Sequence | Inhibition | Folding ability |
|---------|----------|------------|-----------------|
| AAI     | CIPKWNRGCSDKHKW | +          | +               |
| Ins     | CIPKWNLGPPS     | +          | +               |
| A10     | CIPKWNLGPPS     | +          | +               |
| A12     | CIPKWNLGPPS     | +          | +               |
| A14     | CIPKWNLGPPS     | +          | +               |
| A16     | CIPKWNLGPPS     | +          | +               |
| IL12-13 | CIPKWNLGPPS     | +          | +               |
| ChimerA | ACPKWNRCTPG      | +          | +               |
| HWTX-I  | ACKGVDFACTPG     | +          | +               |

*a* Peptides exhibiting \(>1\)% inhibition at 15 \(\mu\)g concentration and 100:1 substrate/enzyme ratio were scored as “+.”  
*b* The folded products (Fig. 1) were purified by reverse phase-HPLC, and their quantity was estimated based on HPLC peak height (taking AAI as 100%). Cases where the quantity of the folded product reached 5% were scored as “+.”  
*c* U, norleucine.

columns for N-terminal sequencing carried out on a MilliGen/Biosearch model 6600 ProSequencer. The released amino acid phenylthiohydantoins were detected simultaneously at 269 and 315 nm.

Peptide Synthesis and Refolding Studies

AAI mutants (Table II) were synthesized manually by solid phase peptide synthesis methods based on FMCOC chemistry and oxidative refolding as described (5). The Huwentoxin I chimera (Table II) was synthesized by FMOC chemistry on a PIONEER peptide synthesizer (PE Biosystems) using FMCOC-L-Leu-polyethylene glycolpolystyrene resin as described (8). The oxidative refolding of the chimera was carried out in the presence of 0.1 mM oxidized glutathione and 1 mM cystine were added immediately before use (the final concentration of the peptide was 20 \(\mu\)g/ml). After 16 h of stirring at 25 °C the reaction was stopped by adjusting the pH to 4.0 with acetic acid. The reaction mixture was analyzed by reverse phase-HPLC as described (5). The folding ability was evaluated from the chromatogram. Typical chromatograms of “well folding” (+) and “poorly folding” (−) samples are shown in Fig. 1A and B, respectively. The peak corresponding to the folded product, indicated by the arrow, was collected and lyophilized for amylase inhibition assay.

NMR Studies

Sample Preparation—Samples were prepared by dissolving AAI powder in 0.5 ml of 20 mM phosphate or acetate buffer (90% \(H_2O, 10\% \text{ D}_2\text{O}\) containing 0.02% Na\text{3} \(\text{PO}_4\) and 0.1 mM EDTA with the final concentration of AAI being 2–4 mM. The pH of the solution was adjusted to 6.5 with 1M HCl and NaOH. Sodium 3-(trimethylsilyl)-2,2,3,3,-tetadeuteropropionate was added as an internal reference for chemical shifts to a final concentration of 200 \(\mu\)M. For the experiments in \(\text{D}_2\text{O},\) AAI samples dissolved in \(\text{H}_2\text{O}\) were lyophilized and redissolved in 99.9% \(\text{D}_2\text{O}\) (Cambridge Isotope Laboratories). The solution was then allowed to stand at room temperature for 24 h before re-lyophilization and a second reconstitution with 99.996% \(\text{D}_2\text{O}.

NMR Spectroscopy—All of the two-dimensional spectra, including COSY, DQF-COSY, TOCSY, and NOESY, were collected at 500 MHz using a Bruker 500 AMX spectrometer at pH 6.5, 300 K. The data points were 512 in the F1 dimension and 2048 in the F2 dimension. Most two-dimensional spectra were recorded by the time proportional phase incrementation method (9). Solvent suppression was carried out by the presaturation method.
Data Processing—SYBYL software (Tripos Inc.) was used on an Indigo 2 Silicon Graphics workstation. All the data were zero-filled to 1000 in the F1 dimension, resulting in a $2048 \times 1024$ ($F_2 \times F_1$) real matrix. The sequential assignment method of Wu thrich (10) allowed identification of all backbone and side chain protons with the exception of a few side chain protons of Ile2, Lys4, Arg7, Lys11, and Glu19. A total of 348 distance constraints were derived from the NOESY spectra (100 and 400 ms in H$_2$O and D$_2$O). 87 long distance constraints, 39 medium distance constraints, 101 sequential constraints, and 121 intraresidue constraints were found. Three distance constraints ($\delta_9$-AAI (Table II) under the same conditions. Column, Supersphere 100 C$_{18}$ 4 mm 4 × 125 column (Merck GmbH); buffer A, 0.1% trifluoroacetic acid in water; buffer B, 0.1% trifluoroacetic acid in acetonitrile; linear gradient from 15 to 40% buffer B in 25 min.

RESULTS

Structural Studies—Previous to the current study, the disulfide topology of AAI was only partially known, i.e. the connectivities of two vicinal Cys residues (17 and 18) were inferred from homology modeling rather than from experimental data. Because this is crucial to the NMR studies, we chose the Edman degradation method combined with phenylthiohydantoin analysis to confirm this disulfide topology at 313 nm. This combined method makes it possible to identify dehydroalanine, the $\beta$-elimination product of phenylthiohydantoin-cystine, which forms when the N-terminal sequencing process reaches a Cys residue disulfide bonded to a sequentially upstream Cys.

Fig. 1. Chromatographic analysis of the folding ability of AAI analogs. A, typical analytical reverse phase- HPLC tracing of a fully reduced synthetic AAI preparation after 16 h of refolding as indicated under “Experimental Procedures.” B, HPLC tracing of a poorly folding peptide, $\Delta\Phi$-AAI (Table II) under the same conditions. Column, Supersphere 100 C$_{18}$ 4 mm 4 × 125 column (Merck GmbH); buffer A, 0.1% trifluoroacetic acid in water; buffer B, 0.1% trifluoroacetic acid in acetonitrile; linear gradient from 15 to 40% buffer B in 25 min.

Fig. 2. Identification of dehydroalanine (indicated as $\Delta$Ser) during Edman sequencing by monitoring the HPLC effluent at 313 nm. Appearance of dehydroalanine in position 18 indicates only that this residue is disulfide-bonded to a sequentially upstream cystein, which, by way of elimination (see “Results”), is Cys1.

Fig. 3. The $\zeta$-trace for the 10 best AAI structures obtained by NMR spectroscopy. The $\zeta$-traces of the 10 best structures (blue) were superimposed, and the root mean square deviations were calculated using the molecular simulation program InsightII. The average root mean square distance was 0.32 Å for the backbone heavy atoms (N, C$_a$, C, and O) and 1.07 Å for all heavy atoms. N and C termini are indicated. The $\zeta$-trace of the final structure is shown in red.

Fig. 4. Ribbon diagram of the three-dimensional structure of AAI. The best structure chosen from the results of PROCHECK (15) calculation is presented as a ribbon diagram produced with the program SETOR (44). The three $\beta$-strands, which form an antiparallel $\beta$-sheet are highlighted by blue arrows, and the three disulfide bridges are indicated in yellow. N and C termini are labeled.
The sequences are shown in Table II. The \(\alpha\)-amylase inhibitory activity was expressed as the difference between the two absorbance values obtained in the absence and in the presence of the peptide inhibitor (i.e., \(A_{546\,\text{nm}}\) control \(A_{546\,\text{nm}}\) peptide). The concentrations of \(\alpha\)-amylase and inhibitor were adjusted so \(A_{546\,\text{nm}}\) was in the optimal range.

\[
\begin{array}{|c|c|c|c|}
\hline
\text{Peptide substrate} & \text{Concentration} & \text{Molecular ratio of peptide to \(\alpha\)-amylase} & \text{Inhibitory activity} \\
\hline
\text{AAI} & 1.4 \times 10^{-5} & 0.14 : 1 & 0.060 \ 100.00 \\
\text{Chimera} & 1.6 \times 10^{-2} & 260 : 1 & 0.076 \ 0.07 \\
\text{Huwentoxin-I} & 1.1 \times 10^{-1} & 1100 : 1 & 0.000 \ 0.00 \\
\hline
\end{array}
\]

\section*{Discussion}

The structure of AAI corresponds to the so-called knottin (3) or inhibitor cystine knot fold (4), which is characterized by three antiparallel \(\beta\)-strands and a disulfide topology of the "abcabc" type, as schematically shown in Fig. 5. The first of the
three β-strands is usually distorted, and it is missing in some members of this family so the core of this motif is often referred to as a short distorted triple-stranded β-sheet. In the solution structure of AAI the first β-strand is in fact less clearly defined than the other two. Members of this family do not show high sequence similarity, and, partly because of their small size, they cannot be easily identified by sequence similarity searches. We previously built a model of AAI based on its homology with the *Trichoderma reesei* CB and squash proteinase inhibitor (2). Although the global features of AAI were correctly captured by this model, the atomic coordinates were only in moderate agreement with the experimental data.

The knottin family contains several distinct subtypes, according to structure-based sequence alignment (Fig. 5). AAI, along with many of the fungal type CBs and some spider toxins, lacks segment 3, so these two central cysteines are vicinal. The CB of *T. reesei* cellobiodylase lacks the disulfide bridge α, whereas ω-conotoxin GVIA has an additional d. Huwentoxin I and other spider toxins form a distinct subgroup characterized by a very short third β-strand. AAI most closely resembles the fungal CB group.

The knottin (3) or inhibitor cystine knot fold (4) was found in various proteins from fungi, plants, spiders, and cone shells. The fact that these proteins fulfill a large variety of biological functions leads us to suggest that this fold may have emerged by convergent evolution. Thus the weak sequence similarities among various knottin proteins may be because of common structural determinants rather than a common evolutionary origin. In fact we have found that certain residues can be replaced without impairing the folding of the molecule, whereas other replacements seem to interfere with the folding. For example we found that segment 1 of AAI can accommodate residue replacements, which is in very good agreement with the homologous region (20). It is interesting, in this respect, that formation of this simple fold is apparently not a one-step process. A two-step mechanism that includes the reshuffling of the first disulfide intermediates was suggested for AAI (5) and also for the related potato proteinase inhibitor (21). It is tempting to speculate that this suggested two-step mechanism may be in fact a condition for the formation of the knottin-structure; quantitative folding studies are now underway.

Members of the knottin family bind to various macromolecular ligands as diverse as cellular receptors, enzymes, or cellulose (for a recent overview see Ref. 20). In the structural alignment in Fig. 5 green displays the residues that are thought to be involved in binding. According to the distribution of these binding residues, knottins can be divided into three broad categories. The proteinase inhibitors from *S. Liang and J. C. Luo, unpublished results.*
Solution Structure of the \(\alpha\)-Amylase Inhibitor AAI

41. Pierpoint, W. S., Tatham, A. S., and Pappin, D. J. C. (1987) *Physiol. Mol. Plant Pathol.* 31, 291–298
42. Singh, N. K., Nelson, D. E., Kuhn, D., Hasegawa, P. M., and Bressan, R. A. (1989) *Plant Physiol. (Bethesda)* 90, 1096–1101
43. Ivengar, R. B., Sits, P., Ouderaa, F. V. d., WEL, H. V. d., Brouwers-Haaven, J. V., Ravensteijn, G., and Wassenaar, P. D. V. (1979) *Eur. J. Biochem.* 96, 193–204
44. Evans, S. V. (1993) *J. Mol. Graph.* 11, 134–138
45. Kraulis, J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J., and Gronenborn, A. M. (1989) *Biochemistry* 28, 7241–7257
46. Davis, J. H., Bradley, E. K., and Basus, V. J. (1993) *Biochemistry* 32, 7396–7405
47. Lampe, R. A. Lo, M. M., Keith, R. A., Horn, M. B., McLane, M. W., Herman, J. L., and Spreen, R. C. (1993) *Biochemistry* 32, 3255–3260
48. Sato, K., Park, N. G., Kohno, T., Maeda, T., Kim, J. I., Kato, R., and Takahashi, M. (1993) *Biochem. Biophys. Res. Commun.* 194, 1292–1296
49. Kim, J. J., Takahashi, M., Oyuma, A., Kohno, T., Kudo, Y., and Sato, K. (1994) *J. Biol. Chem.* 269, 23876–23878
50. Pereira Barbosa, P. J. B., Lozanev, V., Patthy, A., Huber, R., Bode, W., Peng, S., and Strobl, S. (1999) *Structure*, in press