Specific inhibition of insect α-amylases: yellow meal worm α-amylase in complex with the *Amaranth* α-amylase inhibitor at 2.0 Å resolution

Pedro José Barbosa Pereira[1], Valentin Lozanov[2], András Patthy[3], Robert Huber[1], Wolfram Bode[1], Sándor Pongor[2,3] and Stefan Strobl[1]*

Background: α-Amylases constitute a family of enzymes that catalyze the hydrolysis of α-D-(1,4)-glucan linkages in starch and related polysaccharides. The *Amaranth* α-amylase inhibitor (AAI) specifically inhibits α-amylases from insects, but not from mammalian sources. AAI is the smallest proteinaceous α-amylase inhibitor described so far and has no known homologs in the sequence databases. Its mode of inhibition of α-amylases was unknown until now.

Results: The crystal structure of yellow meal worm α-amylase (TMA) in complex with AAI was determined at 2.0 Å resolution. The overall fold of AAI, its three-stranded twisted β sheet and the topology of its disulfide bonds identify it as a knottin-like protein. The inhibitor binds into the active-site groove of TMA, blocking the central four sugar-binding subsites. Residues from two AAI segments target the active-site residues of TMA. A comparison of the TMA–AAI complex with a modeled complex between porcine pancreatic α-amylase (PPA) and AAI identified six hydrogen bonds that can be formed only in the TMA–AAI complex.

Conclusions: The binding of AAI to TMA presents a new inhibition mode for α-amylases. Due to its unique specificity towards insect α-amylases, AAI might represent a valuable tool for protecting crop plants from predatory insects. The close structural homology between AAI and 'knottins' opens new perspectives for the engineering of various novel activities onto the small scaffold of this group of proteins.

Addresses: [1]Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, 82152 Planegg, Germany, [2]International Centre for Genetic Engineering and Biotechnology, Padriciano 99, 34012 Trieste, Italy and [3]Agricultural Biotechnology Centre, Szent-Györgyi A. u. 4, Gödöllö, Hungary.

*Corresponding author.
E-mail: strobl@biochem.mpg.de

Key words: *Amaranthus hypochondriacus*, insect α-amylase inhibitor, knottin, X-ray structure, yellow meal worm

Received: 1 April 1999
Revisions requested: 29 April 1999
Revisions received: 4 May 1999
Accepted: 18 May 1999

Published: 26 August 1999

Structure September 1999, 7:1079–1088
http://biomednet.com/elecref/0969212600701079

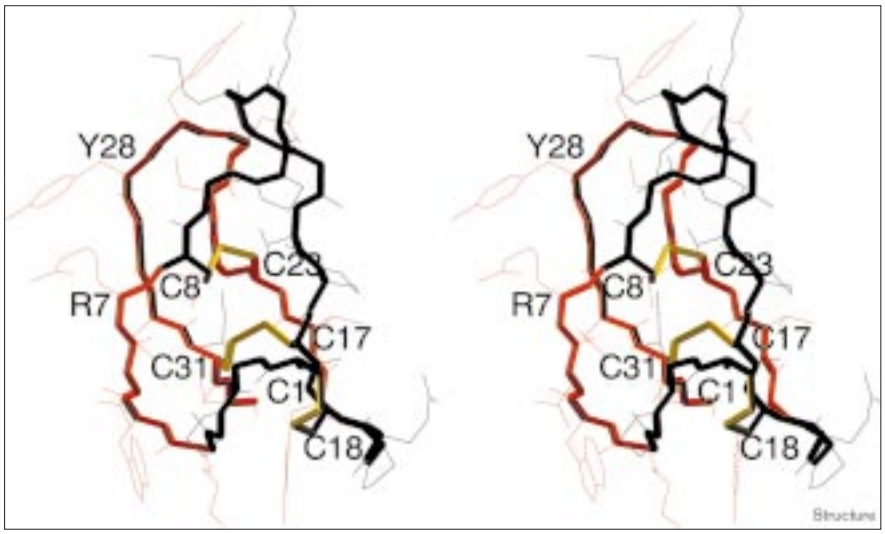
0969-2126/99/$ – see front matter

Introduction

α-Amylases (α-1,4-glucan-4-glucanohydrolase; EC 3.2.1.1) constitute a family of hydrolases that cleave α-D-(1,4)-glucan linkages in starch components, glycogen and various other related carbohydrates. These widely distributed enzymes play an important role in the carbohydrate metabolism of microorganisms, animals and plants. They are the most important digestive enzymes of many insects that feed exclusively on seed products during larval and/or adult life. The α-amylase of the yellow meal worm (TMA), the larval stage of the black meal beetle (*Tenebrio molitor*), is the most extensively studied of these enzymes from insect origin. Its catalytic properties [1,2] and inhibition by several inhibitors *in vitro* [3–9] and *in vivo* [3] have been investigated. We determined the complete amino acid sequence of TMA [2], its three-dimensional structure to 1.64 Å resolution [10] and the structure of its complex with the *Ragi* bifunctional α-amylase/trypsin inhibitor (RBI) at 2.5 Å resolution [11]. Recently, its structure in complex with the bean *Phaseolus vulgaris* α-amylase inhibitor (α-AI) was solved at 3.0 Å resolution [12]. TMA is a monomeric, three-domain (A–C) protein of 471 amino acid residues that contains four disulfide bridges, a structural calcium ion and a chloride ion, which allosterically activates the enzyme. The active site is located in a long, V-shaped substrate-binding cleft at the interface between domains A and B. There, six saccharide units can be accommodated, with the sugar chain being cleaved between the third and fourth bound pyranose rings [10,13]. The catalytic residues are proposed to be Asp185, Glu222, and Asp287 [10]. TMA's general fold is similar to that of the mammalian enzymes; however, important variations in loop segments near the active-site region may be responsible for differences in the susceptibility to inhibitors. Since TMA is biochemically the best characterized insect α-amylase and the only one for which a three-dimensional structure is available, it represents the model enzyme for further studies on the interactions between specific inhibitors and insect α-amylases.

The α-amylase inhibitor from the seeds of *Amaranthus hypochondriacus* (AAI), a Mexican crop plant, is the smallest proteinaceous inhibitor of α-amylases described so far [14,15]. It specifically inhibits insect α-amylases, and is inactive against the mammalian enzymes [14]. AAI consists of 32 amino acid residues and contains three disulfide bridges. The primary sequence is not closely related to any

Figure 1



Stereoview of the AAI model. The hydrophobic hemisphere of the inhibitor (residues 1–3 and 8–20) is shown in black, the hydrophilic hemisphere (residues 4–7 and 21–32) in red. Disulfide bonds are depicted in yellow. Cysteine residues and relevant sidechains are labeled. This figure was made with the program MOLMOL [46].

known protein, but the disulfide topology and residue conservation patterns are similar to those of knottin-like proteins, in particular the squash proteinase inhibitor family, the cellulose-binding domain of cellobiohydrolase, and ω-conotoxins [14]. A three-dimensional model of AAI was built according to this group of proteins, using sidechain replacement and molecular dynamics refinement techniques [14]. Its specificity makes AAI an attractive candidate for conferring pest resistance to transgenic plants. Another interesting feature is its molecular architecture: knottins harbor different biological activities within a small and compact molecular fold, so AAI might be used to design new enzyme inhibitors of medical relevance.

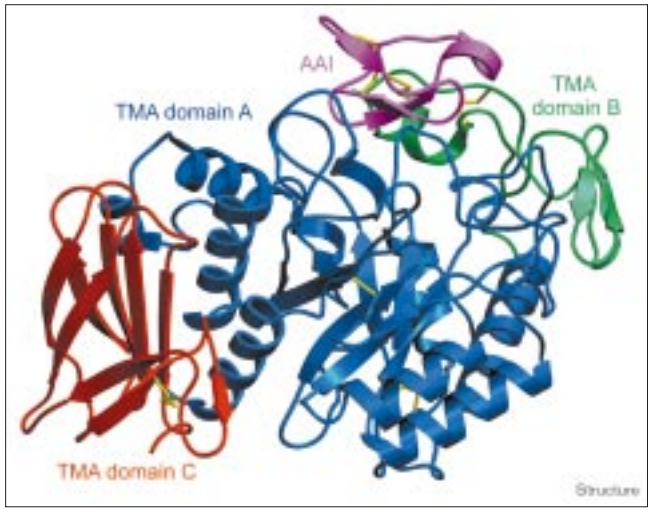
Here we describe the structure of the complex between TMA and AAI at 2.0 Å resolution. It reveals a completely new inhibition mode for proteinaceous α-amylase inhibitors. Two segments of AAI insert into the active-site cleft of TMA and target the catalytic residues. Finally, we analyze the interaction between both proteins with respect to AAI's specificity for insect α-amylases and compare the AAI structure with that of other knottin-like proteins.

Results and discussion

Structure of the *Amaranth* α-amylase inhibitor

AAI exhibits an egg-like shape with approximate dimensions of $26 \times 19 \times 14$ Å (Figure 1). The topology of the 32 amino acid polypeptide chain can basically be described as four strands running sequentially back and forth (Asn6–Gly9, Val15–Glu19, Tyr21–Ser25 and Tyr28–Ser32); however, only three short regular β strands (Asn6–Gly9, Tyr21–Cys23 and Gly29–Cys31), forming a small twisted antiparallel β sheet, can be defined. The only helical secondary structure element, one α-helical turn, is displayed between residues Pro10 and Asp13. There are only 13 intramolecular backbone–backbone hydrogen bonds in AAI, but all parts of the polypeptide chain are uniformly cross-linked via three disulfide bridges (Figures 1,2). The N- and C-terminal ends are localized at the same side of the inhibitor molecule, but are almost 16 Å apart. Both ends are fixed to the molecular core via disulfide bridges, thus creating a rigid and compact structure that is resistant to proteolytic attack by exopeptidases.

Figure 2



Ribbon diagram of the TMA–AAI complex. The three TMA domains, A (residues 1–97 and 160–379), B (residues 98–159) and C (residues 380–471), are depicted in blue, green and red, respectively. AAI is shown in magenta. Disulfide bridges are depicted in yellow. This figure was made with the program SETOR [47].

The disulfide pairing is Cys1–Cys18, Cys8–Cys23 and Cys17–Cys31, that is, the three N-terminal cysteine residues are bound to those of the C-terminal half in an identical order. The conformation of the disulfide bridges is right-handed for the first two and left-handed for the third one. The Cys8–Cys23 and Cys17–Cys31 disulfide bridges are arranged side by side in the core of the molecule (Figure 1). Since all other hydrophobic residues are at least partially accessible to solvent, a real hydrophobic core cannot be defined. On the contrary, the hydrophobic residues are clustered in one hemisphere on the surface of the inhibitor. This region is made up by residues Cys1–Pro3 and Cys8–Pro20. The N terminus and three charged residues (Lys11, Asp13 and Glu19) are evenly distributed over the apolar surface to make it more soluble in an aqueous environment. The other hemisphere, made up by Lys4–Arg7 and Tyr21–Ser32, consists of polar and charged residues. The only hydrophobic residues in this region are Cys23 and Cys31, which connect the hydrophilic half with the hydrophobic half of the molecule, and Trp5, which is involved in blocking the enzyme's active site (see below).

The enzyme–inhibitor interface

The active site of TMA is located in a V-shaped depression at the interface of domains A and B [10,11]. AAI inserts perfectly into this crevice (Figure 2), interacting with residues from both TMA domains that delimit the substrate-binding site. Altogether, 18 inhibitor residues directly interact with 24 residues of the enzyme. Contacts are summarized in Table 1. The total buried surface area amounts to 2085 Å^2 (994 Å^2 for TMA and 1091 Å^2 for AAI) as calculated with the program 'NACCESS' [16]. Although almost half of the inhibitor's accessible surface area is in contact with the enzyme, this is the smallest contact area for an α-amylase–proteinaceous inhibitor complex reported up until now. In other complexes, these contacts range from 2402 Å^2 (TMA–RBI) [11] to 3050 Å^2 (porcine pancreatic α-amylase in complex with α-AI; PPA–α-AI) [17].

TMA's active-site cleft, like that of PPA, provides at least six subsites for the binding of carbohydrate moieties. This was shown by modeling the pseudohexasaccharide inhibitor V-1532 into the active-site cleft of TMA based on the X-ray structure of the PPA–V-1532 complex [11,13] (Figure 3a). In the AAI–TMA complex, carbohydrate-binding subsites 2–5 are completely blocked by residues from two AAI segments (Figure 3b). Tyr28 occupies the position of sugar moiety 2, and Asn6 that of sugar moiety 5. Arg7 resides slightly above subsites 3 and 4. Access to sugar subsites 1 and 6 is sterically prevented by residues Met12 and Tyr27, and Lys4 and Trp5, respectively (Figures 3b,c). Asp287, one of the catalytic residues of TMA, forms a salt bridge directly with Arg7 (AAI). The other two catalytic residues are connected via an intricate

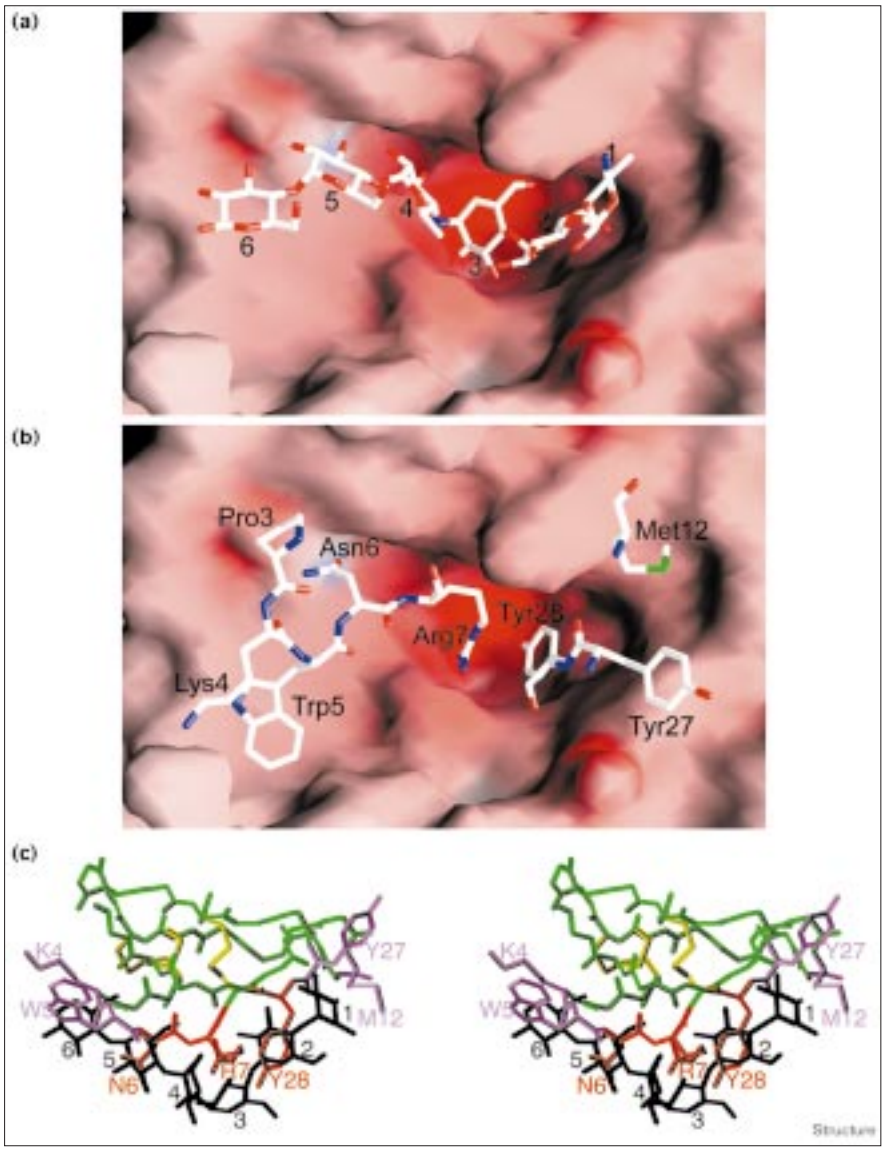
Table 1

Contacts and hydrogen bonds between AAI and TMA.

Residues*			
AAI	TMA	Distance (Å)	Mediated by
Total contacts (distance < 4.0 Å)† between AAI and TMA			
C1	N137, Y139		
P3	Y139, E229		
K4	Q295		
W5	I224, L226, F245, T291, I296		
N6	Y139, K188, E229		
R7	L150, V151, D287		
C8, G9	V151		
M12	G102, V151		
D13	E135, V151		
V15	E135, V136, N137		
T24	R290, T291, G292, N311		
S25	D332		
Y27	W57		
Y28	W56, W57, Y60		
N30	D287, R290, T291		
S32	T291, G292, G293		
Hydrogen bonds (distance < 3.5 Å)† between AAI and TMA			
C1 N	N137 O	2.8	
C1 O	V139 N		wat 667
K4 Nζ	Q295 Oε1	2.7	
W5 O	D287 Oδ2		wat 660
W5 O	D287 Oδ2		wat 707
N6 Oδ1	H189 Nζ		wat 687
N6 Oδ1	K188 Nζ	2.8	
N6 Nδ2	E229 Oε2	3.4	
R7 Nη1	D287 Oδ1	2.8	
R7 Nη1	D287 Oδ2	3.3	
R7 Nη2	D287 Oδ2	3.5	
R7 O	Y139 Oη		wat690
R7 N	H189 Nε2		wat 687
R7 Nη2	E222 Oε1		wat 678
R7 Nη2	E222 Oε2		wat 678
R7 Nη2	D185 Oδ1		wat 678
D13 O	E135 Oε1	2.9‡	
D13 O	E135 Oε2	3.5‡	
D13 Oδ2	V151 N		wat 730
D13 Oδ2	G152 N		wat 732
T24 O	N331 Nδ2	3.2	
S25 Oγ	D332 Oδ1	2.8	
D26 N	D332 Oδ2		wat 777
Y27 O	V151 O		wat 769
Y28 Oη	H99 Nε2		wat 648
Y28 Oη	H286 Nε2		wat 683
Y28 Oη	D185 Oδ2		wat 648
N30 Nδ2	D287 Oδ1		wat 775
N30 Nδ2	D287 O	3.2	
N30 Nδ2	R290 O	3.1	
S32 Oγ	G292 N	3.1	
S32 O	Q295 Oε1		wat 689

*Single-letter amino acid code. †Heavy-atom distances. wat, water molecules. ‡Glu135 (TMA) appears to be protonated. This is conceivable since the crystallization was performed at pH 5.4. Moreover, the hydrogen bond is presumably also of physiological relevance, as in the larval midgut a slightly acidic pH is prevalent [45].

Figure 3

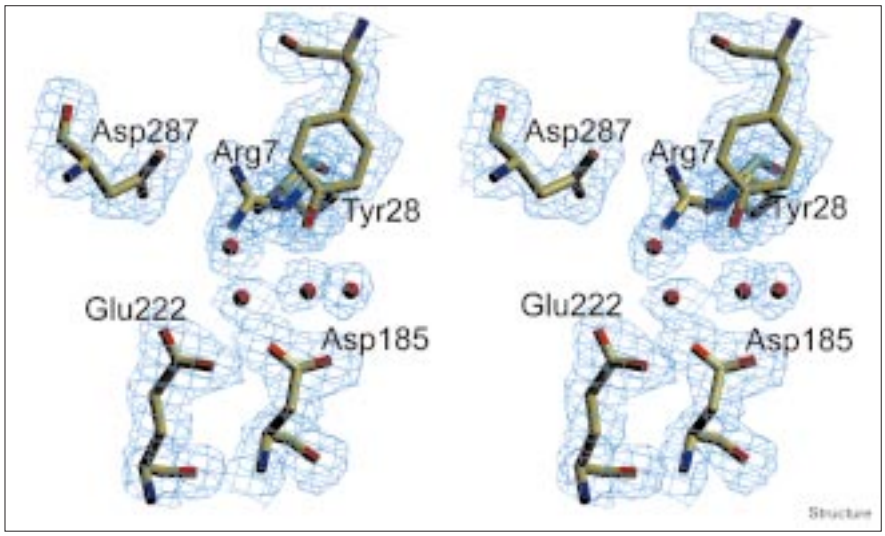


A close up of TMA's active site. (a) A model of the complex between TMA and the hexasaccharide inhibitor V-1532. The model was derived after optimal superposition of the TMA–AAI complex with the PPA–V-1532 complex [13] and removal of the PPA and AAI coordinates. The electrostatic surface potential of TMA is shown (red for partial negative charges, blue for partial positive charges) and the carbohydrate moieties of V-1532 are numbered according to the amylase's substrate-recognition subsites that they are bound to. (b) The same view as in (a), but with selected residues of AAI displayed. (c) Stereoview of the superposition of AAI (as bound to TMA) and the carbohydrate inhibitor V-1532 (as bound to PPA [13]). The backbone atoms of AAI are shown in green. AAI's α-amylase-binding site with selected sidechains is depicted in red and magenta. The carbohydrate units 1–6 of V-1532 are shown in black. Figures (a) and (b) were made with the program GRASP [48] and Figure (c) was prepared with MOLMOL [46].

water-mediated hydrogen-bonding network to the functional groups of inhibitor residues (Figure 4): Glu222 Oε1 and 2 (TMA) are bonded to Arg7 Nη2 (AAI), and Asp185 Oδ2 (TMA) is connected to the hydroxyl group of the Tyr28 (AAI) sidechain and Asp185 Oδ1 to Arg7 Nη2 (AAI). The other TMA residues, which are highly conserved in the active sites of α-amylases and are presumably involved in substrate recognition and/or orientation [10], are also involved in this water-mediated hydrogen-bonding network: His99 Nε2 and His286 Nε2 (both TMA) are linked via water molecules to the hydroxyl group of Tyr28 (AAI), and His189 Nε2 (TMA) interacts with both Asn6 Oδ1 (AAI) and the backbone nitrogen of Arg7 (AAI). Only the ε-amino group of Lys188 (TMA) forms a direct, charged hydrogen bond with Asn6 Oδ1 (AAI). Another nine direct and ten more water-mediated hydrogen bonds between AAI and TMA residues that delimit the active site further stabilize the complex (for comparison, see Table 1). Extended hydrophobic interactions as observed in other complexes of α-amylases and proteinaceous inhibitors cannot be identified [11,17,18]. Thus, the TMA–AAI interface comprises mostly polar interactions, as is also observed for the structure of the complex between barley α-amylase and its endogenous inhibitor [19]. Approximately two thirds of the surface of AAI's hydrophobic hemisphere are solvent exposed. The remaining surface area of this hemisphere interacts exclusively with TMA residues from domain B; however, not by hydrophobic interactions. In contrast, three inhibitor residues from AAI's hydrophilic hemisphere

Figure 4

Stereoview of the electron density of TMA's catalytic residues Asp185, Glu222 and Asp287 and AAI's residues Arg7 and Tyr28. The electron density was calculated with coefficients ($2F_{obs}–F_{calc}$) and contoured at 1.0σ. This figure was made with SETOR [47].



form hydrophobic contacts with TMA residues. The pyrole ring of Trp5 (AAI) interacts with the apolar sidechains of Ile224, Leu226, Phe245, and Ile296 (all TMA), and Tyr28 (AAI) lies with the flat side of its hydroxybenzyl ring in a shallow hydrophobic pocket created by the sidechains of Trp56, Trp57 and Tyr60 (all TMA). The sidechain methylene groups (Cβ–Cδ) of Arg7 (AAI) interact with the parallel-oriented branched sidechains of Leu150 and Val151, which direct the sidechain of Arg7 (AAI) into the center of the active site, where its guanidinium group forms the salt bridge with the carboxylate of Asp287 (TMA; see above).

Altogether, there are only ten direct and seventeen water-mediated hydrogen bonds and one salt bridge established between enzyme and inhibitor. Moreover, the hydrophobic interface between the two proteins is minimal compared with other α-amylase–inhibitor complexes (see above). However, it has to be emphasized that the TMA–AAI complex is characterized by a high complementarity of the interacting surfaces. The inhibitor fills TMA's substrate-binding groove and does not extend considerably beyond this crevice (Figure 2). On the contrary, it confers a conical, lid-like surface to an area that is characterized by a deep cleft in free TMA. Thus, AAI fits perfectly into the active-site cleft of TMA.

A comparison of the molecular models of free TMA and TMA complexed with AAI

Outside of the binding site, no major structural differences between the models of free TMA [10] and TMA complexed with AAI can be detected. Both structures superimpose with root mean square deviation (rmsd) values of 0.48 Å for the backbone atoms and 0.83 Å for all

Figure 5

Stereoview of the superposition of TMA segment His286–Ile296 in the free enzyme (red) and in TMA complexed to AAI (green). Contacting residues from AAI are depicted in blue. This figure was made with MOLMOL [46].

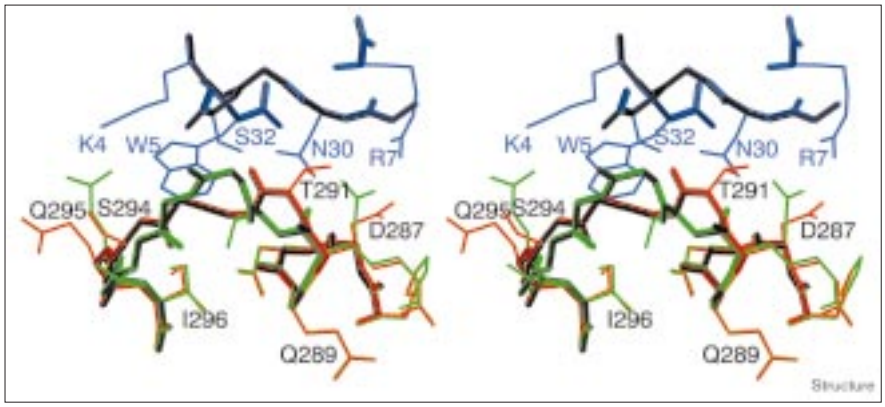


Table 2

Hydrogen bonds between AAI and TMA that cannot be formed in the AAI–PPA complex.

AAI	TMA	Distance* (Å)	Corresponding residue in PPA
Lys4 NZ	Gln295 NE1	2.7	Ser311
Asp13 O	Glu135 OE1†	2.9	Gly147
Thr24 O	Asn331 ND2	3.2	Asp356
Ser25 OG	Asp332 OD1	2.8	Trp357
Asn30 ND2	Arg290 O	3.1	Arg303
Ser32 OG	Gly292 N	3.1	His305

*Heavy-atom distances. †The sidechain carboxyl group of Glu135 (TMA) appears to be protonated. The sidechain is completely defined in the electron-density map (see Table 1).

nonhydrogen atoms. At the interface between enzyme and inhibitor, however, a dramatic change occurs in the backbone conformation of TMA segment Thr291–Ser294. The α helix that is present in free TMA [10] and also in the TMA–RBI complex [11] between residues His286 and Gly292 is shorter by three residues at its C terminus (Figure 5). This conformational change, which is clearly induced by binding of AAI, is accompanied by the formation of two hydrogen bonds between the backbone of TMA and the sidechains of two AAI residues (Arg290 O [TMA] and Asn30 Nδ2 [AAI], and Gly292 N [TMA] and Ser32 Oγ [AAI]). The solvent structure in the central part of the active site is completely changed compared with free TMA, because AAI displaces the water molecules around the catalytic site. Even the water molecule that has been suggested to be the nucleophile attacking the C1 atom of the pyranose ring in sugar-binding subsite 3 [20] is displaced by approximately 1.2 Å. In the structures of free TMA and the TMA–RBI complex it bridges the carboxyl groups of Glu222 and Asp287. In the central part of the active site, only the sidechains of two residues undergo a reorientation upon inhibitor binding: The sidechain of Asp287 is rotated around its Cα–Cβ bond by approximately 65° to form a well-oriented salt bridge with Arg7 (AAI), no longer making contact with the catalytic water, and the benzyl ring of Phe245 moves in the plane of the ring by about 90° around the Cα–Cβ bond in the direction of the sidechain of residue Leu266 to make hydrophobic interactions with the pyrole moiety of AAI residue Trp5. This movement creates enough space in this region of the active site to accommodate an additional water molecule at the position where the Cζ atom of the Phe245 benzyl ring is found in free TMA.

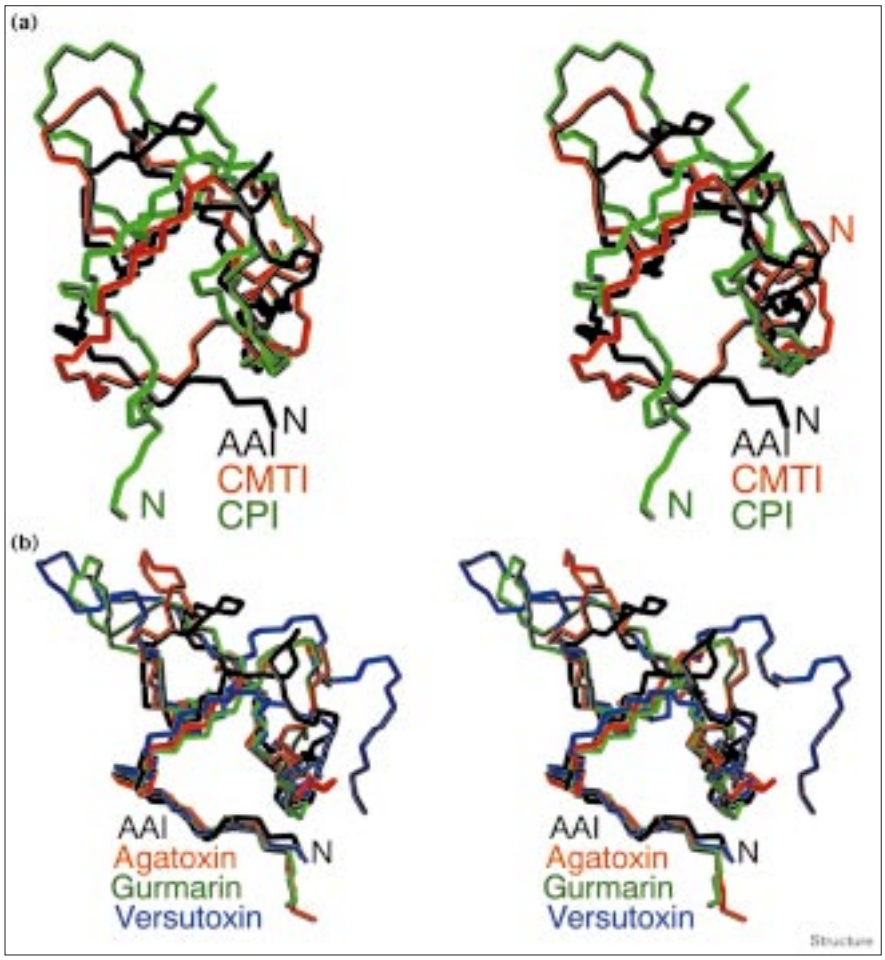
The Ramachandran plots [21] of the (Φ/Ψ) angles demonstrate that the residues in both TMA structures, except for segment Arg290–Ser294, have very similar (Φ/Ψ) angles. This is also true for Met100, which in free TMA is in the additionally allowed region and in the structure of the complex in the generously allowed region (see the Materials and methods section).

AAI specifically inhibits insect α-amylases

Insect and mammalian α-amylases display high homology in their primary and tertiary structures. TMA and porcine pancreatic α-amylase (PPA) share 54% sequence identity [2], and 410 structurally conserved Cα atoms superimpose with an rmsd of 2.2 Å [10]. Nevertheless, there are competitive inhibitors of insect α-amylases that have little or no effect on the mammalian enzymes [4,5,14,22–24] (for a review see [25]). The three-dimensional structure of TMA differs from that of the mammalian α-amylases by the lack of three loops in the vicinity of the active site (residues 140–148, 304–310, and 342–361 in PPA [2,10]). Since the presence and/or absence of these loops is the main structural difference between the α-amylases from both groups of animals, we suggested before that inhibitors exclusively directed against mammalian α-amylases should make crucial contacts to these loops, whereas those targeting solely insect α-amylases should be sterically hindered from binding by at least one of these loops [10].

Previously, three inhibitors of animal α-amylases have been structurally characterized in complex with their target enzymes: the microbial α-amylase inhibitor tendamistat from *Streptomyces tendae* [18] with PPA; the *Ragi* bifunctional α-amylase/trypsin inhibitor with TMA [11]; and the bean α-amylase inhibitor α-AI from *Phaseolus vulgaris* with both α-amylases [12,17]. These three inhibitors are structurally and evolutionarily completely unrelated to one another and bind to α-amylases in entirely different ways. Nevertheless, a comparison of the complexes shows that the inhibitors interact with similar regions of the enzymes in a competitive manner [11]. Recently, the structural determinants for binding of α-AI to PPA and to TMA were compared [12]. It was found that in both complex structures binding of the inhibitor is conferred by almost identical contacts to residues in and around the active-site cleft. Differences only occur in or next to the three additional loops present in PPA but absent in TMA (see above); however, it could be shown that TMA residues neighboring the 'deleted' regions are making up for the missing contacts in an alternative way. Modeling of the tendamistat–TMA complex and the RBI–PPA complex and comparison with the respective experimentally determined structures basically results in a similar conclusion. In this way, all three proteins manage to inhibit both mammalian and insect α-amylases. In contrast to that, AAI was found to specifically inhibit the enzymes from insect sources, whereas it basically does not affect the activity of human or bovine α-amylases [14]. To understand the determinants for its specificity, we modeled AAI into the active site of PPA. In all available PPA models but one, the flexible glycine-rich loop of PPA (residues 304–310) is turned towards the active site, resulting in a

Figure 6



Stereoview of the superposition of the backbones of AAI with other knottin-like proteins. (a) Comparison of AAI (black) with the plant proteinase inhibitors CMTI (*Cucurbita maxima* trypsin inhibitor; red) and CPI (potato carboxypeptidase inhibitor; green). (b) Comparison of AAI (black) with the spider toxins μ-agatoxin-i (red) and versutoxin (blue), and the plant sweet taste-suppressor protein gurmarin (green). This figure was made with MOLMOL [46].

steric clash of the two proteins upon docking. However, in the structure of PPA in complex with α-AI [17], the glycine-rich loop is turned outside due to interactions with the inhibitor. We therefore superimposed the TMA–AAI and the PPA–α-AI complexes, removed TMA and α-AI and evaluated the interactions between PPA and AAI. A comparison with the experimentally determined TMA–AAI complex shows that there are at least six direct hydrogen bonds that can only be formed in the AAI–TMA complex (Table 2). All of the TMA residues participating in these hydrogen bonds are positioned next to or in regions that are structurally dissimilar to PPA; however, four of these hydrogen bonds cannot be formed in a hypothetical PPA–AAI complex because of the enzyme's sidechains. The sidechains of Ser311 and Gly147 (both PPA) are simply not long enough to interact with inhibitor residues Lys4 and Asp13, respectively. And, Asp356 and Trp357 sidechains (both PPA) cannot form hydrogen bonds to Thr24 O and Ser25 Oγ (both AAI), respectively, because their functional groups do not allow this. Two hydrogen bonds cannot be formed between AAI and PPA because of an unfavorable conformation of PPA's backbone: Arg303 O (PPA)–Asn30 Nδ2 (AAI) and His305 N (PPA)–Ser32 Oγ (AAI). We could not, however, identify any steric reasons why AAI should be prevented from binding to PPA by one of the three additional loops in the vicinity of its active site. It rather seems that AAI forms highly specific hydrogen bonds in its complex with TMA. Six out of ten direct hydrogen bonds cannot be formed in an AAI–PPA complex. The lack of extended hydrophobic interactions and the small overall surface interface between enzyme and inhibitor, which would be approximately the same in PPA–AAI complex, are presumably the reasons why no activity against mammalian α-amylases could be detected [14]. Based on these results, we assume that the dissociation constant between AAI and PPA lies far beyond 10^{-5} M whereas that between AAI and TMA should be in the nanomolar range [26].

Structural analogs of AAI

AAI's primary sequence is not closely related to any known protein sequence, but the disulfide topology and residue

conservation patterns are similar to those of members of the squash proteinase inhibitor family, the cellulose-binding domain of cellobiohydrolase, and ω-conotoxins [14]. A model of AAI was built according to the three-dimensional structure of *Ecballium elaterium* trypsin inhibitor (CMTI family) and the cellulose-binding domain of cellobiohydrolase I, using sidechain replacement and molecular dynamics refinement techniques [14]. Although the global fold of AAI was correctly predicted, strong deviations from the experimentally determined X-ray structure (rmsd of 3.7 Å for backbone atoms; data not shown) occur especially in the N-terminal region due to a different orientation of the first disulfide bridge, in the loop with Tyr27 at its tip and in the neighboring α-helical segment (with Met12 at its tip). Since the sidechain of Tyr28 inserts into the active-site crevice of TMA, a rearrangement of these two segments in the free inhibitor seems possible. A more detailed discussion of the quality of the modeled AAI structures should be postponed until the structure of uncomplexed AAI is solved.

The overall structure of AAI, its three antiparallel β strands forming a twisted β sheet, the disulfide-bonding pattern and the topological location of the second and third cystine bridges clearly identify it as a member of the so-called knottins [27], a group of small proteins that bind to a variety of macromolecular ligands such as cellulose, enzymes and cellular receptors [28]. A structural alignment with the squash seed trypsin inhibitor (CMTI) [29] and potato carboxypeptidase inhibitor (CPI) [30] shows a striking structural similarity of AAI with these plant proteinase inhibitors (Figure 6a). Since there are several independent examples of bifunctional α-amylase/proteinase inhibitors in the plant kingdom [25], it also seems conceivable that inhibitors of both classes of digestive enzymes developed on the same homologous scaffold from a common precursor protein. Members of the cereal inhibitor superfamily, for example, share an identical folding pattern (four helices arranged in a simple up-down motif) and disulfide topology with the plant nonspecific lipid transfer proteins, the 2S-storage proteins [31] and the hydrophobic protein from soybean [32], thus demonstrating that the fold can be preserved while the function of the protein is modulated. However, a DALI search [33] on AAI confirmed that the structurally most similar proteins are μ-agatoxin-i from the funnel-web spider (PDB entry 1eit [34]), gurmarin from *Gymnema sylvestre* (PDB entry 1gur [35]) and versutoxin from the blue mountains funnel-web spider (PDB entry 1vtx [36]), typical members of the knottin family (Figure 6b) for which an evolutionary relationship with AAI can be excluded. Obviously many small and probably completely unrelated proteins rely on the same overall folding pattern and a similar disulfide geometry to maintain their three-dimensional structures. Since AAI does not share any other important motifs (conserved, structurally important non-cysteine residues, or the same spacer length between cysteines) with CMTI and CPI besides the three disulfide bonds and the small twisted antiparallel β sheet, a possible common evolutionary root of these inhibitors remains mere speculation.

The knottins perform a variety of different tasks [28] but share an almost identical scaffold. This could be used to design proteins with new activities of medical and/or industrial importance, either by genetic engineering methods or by peptide synthesis and refolding. Preliminary experiments to change the specificity of AAI towards a proteinase inhibitor yielded promising results (VL and SP, unpublished results). As successful synthesis and refolding [15] has been demonstrated for AAI, and the three-dimensional structure and the mode of interaction with its target is known, it represents an ideal candidate for experiments in this direction.

Biological implications

Insects like *Tenebrio molitor*, which are widespread pests of plants, are responsible for enormous losses of foodstuffs by damaging grain and stored crops. Carbohydrate in the form of starch forms the bulk of their diet. α-Amylases perform the first step of enzymatic starch degradation. Knowledge about the inhibition of these important digestive enzymes of insect pests is therefore of paramount interest for phytopharmacy.

Here, we report the structure of yellow meal worm α-amylase (TMA) in complex with the α-amylase inhibitor (AAI) from the Mexican crop plant *Amaranthus hypochondriacus*. With 32 amino acids, AAI is the smallest α-amylase inhibitor described so far. Residues from two AAI segments target the active site of TMA, resulting in a completely new mechanism of inhibition of the enzyme.

Due to its specificity against insect-derived α-amylases, AAI might present a valuable tool for the protection of crop plants from predatory insects. Several ways of achieving this can be envisaged, among them the genetic manipulation of the plants of interest, or alternatively the use of AAI or of engineered derivatives of this inhibitor as 'biologically safe' insecticides. Moreover, and again due to its selective mode of inhibition, AAI represents a promising starting point for the structure-based rational design of specific synthetic inhibitors of α-amylases, with a wide potential use in the protection of economically important crops.

The knottin fold is present in a wide variety of small proteins, displaying very diverse functions. Although they have low (if any) sequence homology, these proteins share a similar overall structure, relying on subtle chemical differences to accomplish their specific tasks. AAI, as a

member of this group of proteins, could be used as a prototype to further investigate their mode of action. This could be accomplished by engineering AAI in order to develop novel activities for this efficient insect α-amylase inhibitor.

Materials and methods

Protein purification and crystallization

TMA and AAI were prepared as described before [2,15]. Prior to crystallization, concentrated samples of both proteins were mixed, incubated for 30 minutes at room temperature and subjected to size-exclusion chromatography on a Superdex 75 HighLoad 26/60 column (Pharmacia), equilibrated with 5 mM sodium acetate (pH 5.4), 0.1 mM $CaCl_2$. Fractions containing the AAI–TMA complex were pooled and concentrated to approximately 10 mg/ml. Crystals were grown at 23°C from 10 μl droplets using the hanging-drop vapor-diffusion method. The droplets consisted of 7.5 μl protein solution and 2.5 μl precipitant (12% [w/v] PEG 1000, 12% [w/v] PEG 8000) and were equilibrated against 500 μl of precipitant solution. Crystals of space group $P6_1$, with cell constants a = b = 119.25 Å and c = 64.44 Å, appeared after one day and grew to a maximum size of approximately $150 \times 150 \times 750$ μm.

Data collection, structure solution and refinement

Crystals were mounted in thin-walled glass capillaries. X-ray diffraction data from a single crystal were recorded at 16°C on an image-plate detector (MAR Research, Hamburg, Germany) attached to a Rigaku-Denki rotating Cu-anode generator providing graphite-monochromatized CuKα radiation. Ninety 1°-rotation images were collected without any significant decay in the crystal diffraction properties. Data were processed using the program DENZO [37] and routines from the CCP4 suite [38]. The three-dimensional structure was solved by molecular replacement using the coordinates of unliganded TMA [10]. The program AMoRe [39] was used for rotational and translational searches, finding one complex per asymmetric unit, which is in good agreement with the value for the Matthews coefficient previously calculated (2.25 $Å^3$/Da, corresponding to 45% solvent content). The best solution (43.6; 106.2; 195.8; 0.4078; 0.2049; 0; α, β, γ are given in Eulerian angles; *x, y* and *z* are fractional cell coordinates) had a correlation coefficient of 71.8% and a crystallographic R factor of 29.1% for data between 10.0 and 4.0 Å resolution (the second best solution had values of 37.6% for the correlation coefficient and 42.0% for the R factor). After positional refinement with X-PLOR [40], the calculated $(2F_{obs} - F_{calc})$ and $(F_{obs} - F_{calc})$ electron-density maps permitted modification of the search model and building of the inhibitor using the program TurboFRODO [41]. The model was refined using the Engh & Huber [42] target values and individual temperature factors.

The final model of the TMA–AAI complex was refined to a crystallographic R factor of 16.1% (free R factor 19.0%) for 32,857 unique reflections between 8.0 and 2.0 Å resolution. The statistics of the data collection, processing and structure determination are given in Table 3. The final model contains all 471 TMA residues and all 32 AAI residues. One calcium cation and one chloride anion were assigned at positions homologous to those in unliganded TMA based on strong electron-density peaks and their coordination geometry. A patch residue with appropriate topology and parameter values was set up for the N-terminal pyroglutamate of TMA for refinement with X-PLOR [40]. Water molecules (273) were additionally introduced at stereochemically reasonable positions.

Quality of the final model

The rmsds from standard values were determined to be 0.007 Å for bond lengths and 1.778° for angles (Table 3). The quality of the model was assessed with the program PROCHECK [43]. A Ramachandran plot [21] of the Φ/Ψ angles indicates that 87% of the amino acids are in the most favoured regions and 12% are in additionally allowed regions. Met100 and Asp390 (both TMA) are in the generously allowed regions, and Ser12 (TMA) is localized in the disallowed region. However, the sidechains of these residues are well defined by electron density, like those of most other residues. Exceptions are the surface-exposed sidechains of TMA residues Asp213 (disordered Oδ1 and Oδ2), Asp350, Asn352, Asp446 (disordered beyond Cα), Glu235, Glu378, Asp451, Asp459 (disordered beyond Cβ) and Gln381 (disordered beyond Cγ), and of AAI residue Lys11 (disordered beyond Cβ). None of these residues is involved in enzyme–inhibitor interactions.

Table 3

Summary of data collection and refinement.

Space group	$P6_1$
Cell constraints (Å)	a = b =119.25; c = 64.78; α = β = 90°; γ = 120°
Number of measurements	279,403
Number of unique reflections	33,598
Average multiplicity	8.3
R_{merge}* (%) 20.0–2.0/2.05–2.00† Å	13.5/33.9
Completeness (%) 20.0–2.0 /2.05–2.00 Å	94.4 /68.3
Resolution range for refinement (Å)	8.0–2.0
Reflections used for refinement	32,857
Working set	31,194
Test set	1,663
R_{cryst}‡	16.1
R_{free}§	19.0
Rmsd bond lengths (Å)	0.007
Rmsd bond angles (°)	1.778
Rmsd dihedral angles (°)	23.476
Rmsd improper angles (°)	1.303
Rmsd bonded B factors ($Å^2$)	1.951
Number of nonhydrogen protein atoms¶	3606/246
B factors for protein atoms ($Å^2$)	
Average¶	18.202/18.526
Maximum¶	60.130/57.720
B factors for solvent atoms ($Å^2$)	
Average	30.152
Maximum	65.880
Solvent molecules	273
Inorganic ions	
Calcium (CA W 601)	1 (B factor 13.12)
Chloride (CL W 602)	1 (B factor 24.97)

*$R_{merge} = \Sigma_{h,k,l}\Sigma_i |I(h,k,l) - I(h,k,l)| / \Sigma_{h,k,l}\Sigma_i |I_i(h,k,l)|$, where (h,k,l) are the unique reflections, $I_i(h,k,l)$ is the i^{th} measurement intensity of reflection h,k,l and |I(h,k,l)| is the mean of all measurements h,k,l. †Last shell. ‡$R_{cryst} = \Sigma_{h,k,l} ||F_{obs}(h,k,l)| - |F_{calc}(h,k,l)|| / \Sigma_{h,k,l} |F_{obs}(h,k,l)|$, where F_{obs} (h,k,l) and F_{calc} (h,k,l) are the observed and calculated structure factors, respectively. §R_{free} is the R value calculated for a subset of 5% of randomly selected reflections not used in the refinement. ¶TMA/AAI.

Accession numbers

The atomic coordinates of the TMA–AAI complex have been deposited with the Brookhaven Protein Data Bank with reference number 1clv. They will be released with a delay of one year, but are available from the authors on request in the meantime.

Note added in proof

Recently, the NMR structure of the *Amaranth* α-amylase inhibitor, PDB accession code 1qfd, has been determined [44].

Acknowledgements

This work was supported by scholarship Praxis XXI/BD/9782/96 to PJBP from Fundação para a Ciência e a Tecnologia, Portugal. PJBP is a 'Programa Gulbenkian de Doutoramento em Biologia e Medicina' fellow. VL was the recipient of an ICGEB fellowship. The financial support by the SFB469 and the Biotech program BIO4-CT98-0418 of the European Union (WB and SS) is gratefully acknowledged.

References

1. Buonocore, V., Poerio, E., Silano, V. & Tomasi, M. (1976). Physical and catalytic properties of α-amylase from *Tenebrio molitor* L. larvae. *Biochem. J.* 153, 621-625.
2. Strobl, S., Gomis-Rüth, F.X., Maskos, K., Frank, G., Huber, R. & Glockshuber, R. (1997). The α-amylase from the yellow meal worm: complete primary structure, crystallization and preliminary X-ray analysis. *FEBS Lett.* 409, 109-114.
3. Applebaum, S.W. (1964). The action pattern and physiological role of *Tenebrio* larval amylase. *J. Ins. Physiol.* 10, 897-906.
4. Shainkin, R. & Birk, Y. (1970). Alpha-amylase inhibitors from wheat. Isolation and characterization. *Biochim. Biophys. Acta* 221, 502-513.
5. Silano, V., Pocchiari, F. & Kasarda, D.D. (1973). Physical characterization of α-amylase inhibitors from wheat. *Biochim. Biophys. Acta* 317, 139-148.
6. Buonocore, V. & Poerio, E. (1976). Interaction of *Tenebrio molitor* L. α–amylase with a wheat flour protein inhibitor. *FEBS Lett.* 67, 202-206.
7. Silano, V. & Zahnley, J.C. (1978). Association of *Tenebrio molitor* L. α-amylase with two protein inhibitors – one monomeric, one dimeric – from wheat flour. Differential scanning calorimetric comparison of heat stabilities. *Biochim. Biophys. Acta* 533, 181-185.
8. Buonocore, V., Gramenzi, F., Pace, W., Petrucci, T., Poerio, E. & Silano, V. (1980). Interaction of wheat monomeric and dimeric protein inhibitors with α-amylase from yellow mealworm (*Tenebrio molitor* L. larva). *Biochem. J.* 187, 637-645.
9. Garcia-Maroto, F., Carbonero, P. & Garcia-Olmedo, F. (1991). Site-directed mutagenesis and expression in *Escherichia coli* of WMAI-1, a wheat monomeric inhibitor of insect α-amylase. *Plant Mol. Biol.* 17, 1005-1011.
10. Strobl, S. *et al.*, & Glockshuber, R. (1998). Crystal structure of yellow meal worm α-amylase at 1.64 Å resolution. *J. Mol. Biol.* 278, 617-628.
11. Strobl, S., Maskos, K., Wiegand, G., Huber, R., Gomis-Ruth, F.X. & Glockshuber, R. (1998). A novel strategy for inhibition of α-amylases: yellow meal worm α-amylase in complex with the *Ragi* bifunctional inhibitor at 2.5 Å resolution. *Structure* 6, 911-921.
12. Nahoum, V. *et al.*, & Payan, F. (1999). A plant seed inhibitor of two classes of α-amylases: X-ray analysis of *Tenebrio molitor* larvae α-amylase in complex with the bean *Phaseolus vulgaris* inhibitor. *Acta Crystallogr. D* 55, 360-362.
13. Machius, M., Vertesy, L., Huber, R. & Wiegand, G. (1996). Carbohydrate and protein-based inhibitors of porcine pancreatic α-amylase: structure analysis and comparison of their binding characteristics. *J. Mol. Biol.* 260, 409-421.
14. Chagolla-López, A., Blanco-Labra, A., Patthy, A., Sanchez, R. & Pongor, S. (1994). A novel α-amylase inhibitor from amaranth (*Amaranthus hypocondriacus*) seeds. *J. Biol. Chem.* 269, 23675-23680.
15. Lozanov, V., Guarnaccia, C., Patthy, A., Foti, S. & Pongor, S. (1997). Synthesis and cystine/cysteine-catalyzed oxidative folding of the amaranth α-amylase inhibitor. *J. Pept. Res.* 50, 65-72.
16. Hubbard, S.J. & Thornton, J.M. (1993), *NACCESS*, Department of Biochemistry and Molecular Biology, University College, London.
17. Bompard-Gilles, C., Rousseau, P., Rouge, P. & Payan, F. (1996). Substrate mimicry in the active center of a mammalian α-amylase: structural analysis of an enzyme–inhibitor complex. *Structure* 4, 1441-1452.
18. Wiegand, G., Epp, O. & Huber, R. (1995). The crystal structure of porcine pancreatic α-amylase in complex with the microbial inhibitor Tendamistat. *J. Mol. Biol.* 247, 99-110.
19. Vallée, F. *et al.*, & Haser, R. (1998). Barley α-amylase bound to its endogenous protein inhibitor BASI: crystal structure of the complex at 1.9 Å resolution. *Structure* 6, 649-659.
20. Mazur, A.K., Haser, R. & Payan, F. (1994). The catalytic mechanism of α-amylases based upon enzyme crystal structures and model building calculations. *Biochem. Biophys. Res. Commun.* 204, 297-302.
21. Ramachandran, G.N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advan. Protein Chem.* 23, 283-437.
22. Silano, V. *et al.*, & Valfre, F. (1975). Inhibition of amylases from different origins by albumins from the wheat kernel. *Biochim. Biophys. Acta* 391, 170-178.
23. Buonocore, V., Petrucci, T. & Silano, V. (1977). Wheat protein inhibitors of α-amylase. *Phytochemistry* 16, 811-820.
24. Gutiérrez, C., Sanchez Monge, R., Gomez, L., Ruiz Tapiador, M., Castanera, P. & Salcedo, G. (1990). Alpha-amylase activities of agricultural insect pests are specifically affected by different inhibitor preparations from wheat and barley endosperms. *Plant Sci.* 72, 37-44.
25. Richardson, M. (1990) In *Methods in Plant Biochemistry* (Rogers, L., ed.), Vol. 5, pp. 261-307 Academic Press, London.
26. Bieth, J. (1974) In *Bayer-Symposium 5: Proteinase Inhibitors* (Fritz, H., Tschesche, H., Greene, L.J. and Truscheit, E., eds.), pp. 463-469, Springer-Verlag, Berlin, FRG.
27. Lin, S.L. & Nussinov, R. (1995). A disulphide-reinforced structural scaffold shared by small proteins with diverse functions. *Nat. Struct. Biol.* 2, 835-837.
28. Smith, G.P., Patel, S.U., Windass, J.D., Thornton, J.M., Winter, G. & Griffiths, A.D. (1998). Small binding proteins selected from a combinatorial repertoire of knottins displayed on phage. *J. Mol. Biol.* 277, 317-32.
29. Bode, W., Greyling, H.J., Huber, R., Otlewski, J. & Wilusz, T. (1989). The refined 2.0 Å X-ray crystal structure of the complex formed between bovine β-trypsin and CMTI-I, a trypsin inhibitor from squash seeds (*Cucurbita maxima*). Topological similarity of the squash seed inhibitors with the carboxypeptidase A inhibitor from potatoes. *FEBS Lett.* 242, 285-292.
30. Rees, D.C. & Lipscomb, W.N. (1982). Refined crystal structure of the potato inhibitor complex of carboxypeptidase A at 2.5 Å resolution. *J. Mol. Biol.* 160, 475-498.
31. Behnke, C.A. *et al.*, & Teller, D.C. (1998). Structural determinants of the bifunctional corn Hageman factor inhibitor: X-ray crystal structure at 1.95 Å resolution. *Biochemistry* 37, 15277-15288.
32. Oda, Y., Matsunaga, T., Fukuyama, K., Miyazaki, T. & Morimoto, T. (1997). Tertiary and quaternary structures of 0.19 α-amylase inhibitor from wheat kernel determined by X-ray analysis at 2.06 Å resolution. *Biochemistry* 36, 13503-13511.
33. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* 233, 123-138.
34. Omecinsky, D.O., Holub, K.E., Adams, M.E. & Reily, M.D. (1996). Three-dimensional structure analysis of μ-agatoxins: further evidence for common motifs among neurotoxins with diverse ion channel specificities. *Biochemistry* 35, 2836-2844.
35. Arai, K. *et al.*, & Akasaka, K. (1995). Three-dimensional structure of gurmarin, a sweet taste-suppressing polypeptide. *J. Biomol. NMR* 5, 297-305.
36. Fletcher, J.I., Chapman, B.E., Mackay, J.P., Howden, M.E. & King, G.F. (1997). The structure of versutoxin (δ-atracotoxin-Hv1) provides insights into the binding of site 3 neurotoxins to the voltage-gated sodium channel. *Structure* 5, 1525-1535.
37. Otwinowski, Z. & Minor, W. (1993) In *DENZO: A Film Processing for Macromolecular Crystallography*, Yale University, New Haven.
38. Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* 50, 760-763.
39. Navaza, J. (1994). An automated package for molecular replacement. *Acta Cryst. A* 50, 157-163.
40. Brünger, A.T. (1991). Crystallographic phasing and refinement of macromolecules. *Curr. Opin. Struct. Biol.* 1, 1016-1022.
41. Roussel, A. & Cambilleau, C. (1989) In *TurboFRODO in Sillicon Graphics Geometry*, Partners Directory, Sillicon Graphics, Mountain View, California.
42. Engh, R. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst. A* 47, 392-400.
43. Laskowski, R., MacArthur, M., Hutchinson, E. & Thorton, J. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283-291.
44. Lu, S., *et al.*, & Pongor, S. (1999). Solution structure of the major α-amylase inhibitor of the crop plant Amaranth. *J. Biol. Chem.* 274, 20473-20478.
45. Srivastava, U.S. & Srivastava, P.D. (1961). On the hydrogen-ion concentration in the alimentary canal of the coleoptera. *Beitr. Entemol.* 11, 15-20.
46. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 51-55.
47. Evans, S.V. (1990). SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graphics* 11, 134-138.
48. Nicholls, A., Bharadwaj, R. & Honig, B. (1993). GRASP – graphical representation and analysis of surface properties. *Biophys. J.* 64, A166.