A Multigene Family of Trypsin/α-Amylase Inhibitors from Cereals

PILAR CARBONERO and FRANCISCO GARCÍA-OLMEDO
Laboratorio de Bioquímica y Biología Molecular, Departamento de Biotecnología - UPM, E.T.S. Ingenieros Agrónomos, E-28040 Madrid, Spain

1. CHARACTERISTICS.

1.1 Distribution and solubility.

Plant proteinaceous inhibitors of hydrolases from heterologous systems (fungi, insects, vertebrates, etc.) have been extensively studied. Recent advances in the molecular biology of these inhibitors have greatly increased our knowledge of their structure and *in vitro* properties, and have allowed previously unsuspected relationships between functionally different proteins to be established.

In wheat and barley, a substantial fraction of the total endosperm protein content is represented by toxins and inhibitors that are active towards heterologous systems. In these cereals, a single protein family of trypsin/α-amylase inhibitors is prominently represented among albumins and globulins from endosperm. More than 20 different members from this family have been characterized (reviewed by Carbonero *et al.*, 1993; García-Olmedo *et al.*, 1987; 1992). Their apparent molecular weights are in the 12-16,000 range, and some of them can be selectively extracted with chloroform/methanol mixtures and have therefore been designated CM-proteins. The α-amylase inhibitors can be classified according to their degree of aggregation into monomeric, dimeric, and tetrameric forms. The trypsin inhibitors are monomeric. No members of this family have been found in tissues other than endosperm, although this aspect deserve further investigation. During kernel development, their synthesis precedes that of the main storage proteins and they are rapidly degraded upon germination. These inhibitors have been found not only in wheat and barley but also in other species of the Poaceae (Gramineae) such as rye, rice, maize and finger-millet (Carbonero *et al.*, 1993; García-Olmedo *et al.*, 1987; 1992).

1.2 Amino acid sequences

Amino acid sequences of members of this inhibitor family from wheat, barley and other cereals - directly determined or deduced from nucleotide sequences of cDNA clones - have been aligned in Figure 1. The sequences have been organized into three domains (A, B, C) and grouped into ten subfamilies according to sequence similarity and *in vitro* activity. In addition to the homology relationships summarized in Figure 1, weaker and more elusive relationships have been proposed between this family and the 2S storage proteins from dicots, with dispersed sulphur-rich domains from cereal prolamins and even with the Kazal secretory trypsin inhibitor from bovine pancreas (reviewed in García-Olmedo *et al.*,1992).

Barley trypsin inhibitor BTI-CMe, one of the best characterized members of the family (Odani *et al.*, 1983; Rodriguez-Palenzuela *et al.*, 1989; Royo *et al.*, 1996; Shewry *et al.*, 1984), belongs to the same subfamily as the trypsin inhibitors from rye (RTI), maize (MTI) and ragi (RBI) (Campos and Richardson, 1983; Lyons *et al.*, 1987; Mahoney *et al.*, 1984; Wen *et al.*, 1992). The wheat homologue has not yet been isolated, although a related cDNA (pCMx) has been characterized (Sanchez de la Hoz *et al.*, 1994). The reactive site of BTI-CMe is the motif glycine-proline-arginine-leucine (GPRL) that is located at the right-hand border of the A domain, a region that is quite variable throughout the family. This same reactive site appears in RTI, MTI and RBI. Two related members, represented by cDNA clones pUP23 from barley (Lazaro *et al.*, 1988a) and its wheat homoeologue pUP-88 have the variant sequence proline-serine-arginine-proline in the same position (Carbonero *et al.*, 1993). However, *in vitro* inhibition tests of the native proteins would be needed to confirm that these proteins are not trypsin inhibitors.

The monomeric α-amylase inhibitor from barley, BMAI-1, whose amino acid sequence has been deduced from its cDNA, is glycosylated, a feature that is related to its allergenic properties (Barber *et al.*, 1989; Mena *et al.*, 1992), as will be discussed below. An homologous protein from rice, RAP, is also a major allergen and its cDNA and genomic DNA have been characterized (Adachi *et al.*, 1993; Izumi *et al.*, 1992). This protein has an extension of six amino acids (aspartate-histidine-histidine-glutamine-valine-tyrosine) with respect to the *N*-terminal serine of the barley BMAI-1 inhibitor (Izumi *et al.*, 1992; Mena *et al.*, 1992). Two homologous monomeric α-amylase inhibitors have been purified from wheat and the cDNA from one of them (WMAI-1) has been characterized (García-Maroto *et al.*, 1991; Gomez *et al.*, 1991; Kashlan and Richardson, 1981).

The amino acid sequences of the homodimeric α-amylase inhibitors encoded by genes from the B and D genomes of wheat are close to each other, but that of the barley homodimeric inhibitor, BDAI-1, is closer to those of the

wheat monomeric inhibitors than to those of either the barley monomeric or the wheat dimeric inhibitors (Carbonero *et al.*, 1993; Lazaro *et al.*, 1988b; Maeda *et al.*, 1985).

Fig. 1. Alignment of amino acid sequences of cereal inhibitors of α-amylase/trypsin. Sequences are divided into A, B and C domains and have been grouped according to similarity and *in vitro* activities. The reactive sites of trypsin inhibitors (proline-arginine-leucine) at the end of the A domain are underlined indicated with a wavy line and putative *N*-glycosylation sites (asparragine-X-serine/threonine) in pUP23, BMAI-1, BDA-1, WTAI-CM16, WTAI-CM17 and BTAI-CMb are indicated with a double underline. Only partial *N*-terminal sequences are available for WMAI-2 and WDAI-3.

```
A
BTI-CME    --FGDSCAPGDALPH NPLRACRTYVVSQIC HQGPRLLTSD----- ----------
RTI        -SVGGQCVPGLAMPH NPLGACRTYVVSQIC HVGPRLFTWD----- ----------
RBI        -SVGTSCIPGMAIPH NPLDSCRWYVSTRTC GVGPRLATQE----- ----------
MTI        -SAGTSCVPGWAIPH NPLPSCRWYVTSRTC GIGPRLPWPEGRLE- ----------

PUP-23     -SVKDECQLGVDFPH NPLATCHTYVIKRVC GRGPSRPMLV----- ----------
PUP-88     -SVEDECQPGVAFPH NALATCHTYVIKRVC GRGPSRPMLV----- ----------

BTI-CMc    -TSIYTCYEGMGLPV NPLQGCRFYVASQTC GAVPLLPIEV----- ----------

BMAI-1     -SPGEWCWPGMGYPV YPFPRCRALVKSQ-C AG-GQVVESIQ---- ----------
RAP        -SPGEQCRPGISYPT YSLPQCRTLVRRQ-C VGRGASAADEQV--- ----------

BDAI-1     SGPWMWCDPEMGHKV SPLTRCRALVKLE-C VG-----NRVPEDVL ----------

WMAI-1     SGPWSWCNPATGYKV SALTGCRAMVKLQ-C VGSQVPEAVL----- ----------
WMAI-2     SGPWMWCDPAMGYRV SPLTGCRAMVKLQ-C VGSQVPEA------- ----------

WDAI-1     SGPWM-CYPGQAFQV PALPGCRPLLKLQ-C NGSQVPEAVL----- ----------
WDAI-2     SGPWM-CYPGQAFQV PALPACRPLLRLQ-C NGSQVPEAVL----- ----------
WDAI-3     SGPWM-CYPGYAFKV PALPGCRPVLLLQ-C NGSQVPEAVL----- ----------

WTAI-CM1   --TGPYCYAGMGLPI NPLEGCREYVASQTC GIS-ISGSAVSTEPG NT---------
WTAI-CM2   --TGPYCYPGMGLPS NPLEGCREYVAQQTC GVGIIVGSPVSTEPG NT---------
BTAI-CMa   --TGQYCYAGMGLPS NPLEGCREYVAQQTC GVT-IAGSPVSSEPG DT---------

WTAI-CM16  -IGNEDCTPWMSTLI TPLPSCRDYVEQQAC RIETPGS-------- ----------
WTAI-CM17  ---NEDCTPWTSTLI TPLPSCRNYVEEQAC RIEMPGPPYL----- ----------
BTAI-CMb   -VGSEDCTPWTATPI TPLPSCRDYVEQQAC RIETPGPPYL----- ----------
```

```
WTAI-CM3   ---SGSCVPGVAFRT NLLPHCRDYVLQQTC TFTPGSKLPEWMTSA S-IYSPGKPYL
BTAI-CMd   AAAATDCSPGVAFPT NLLGHCRDYVLQQTC AVLTPGSKLPEWMTS AELNYPGQPYL
```

B

```
BTI-CME    MKRRCCDELSAIP- AYCRCEALRIIMQGV VTWQGA--------F EGAYFK----
RTI        MKRRCCDELLAIP- AYCRCEALRILMDGV VTQQGV--------F EGGYLK----
RBI        MKARCCRQLEAIP- AYCRCEAVRILMDGV VTSSGQ--------H EGRLLQ----
MTI        LKRRCCRELADIP- AYCRCTALSILMDGA IPP-GP--------- DAQLE-----

PUP-23     -KERCCRELAAVP- DHCRCEALRILMDGV RTPEG-RWEGRLG-- ----------
PUP-88     -KERCCRELAVVP- DYCRCEALRVLMDGV RAEEGHVVEGRLG-- ----------

BTI-CMc    MKDWCCRELAGISS N-CRCEGLRVFIDRA FPPSQSQ--GAPPQL PPL-------

BMAI-1     --KDCCRQIAAIGD EWCICGALGSMRGSM YKELGVA-------- LADDKATVAE
RAP        -WQDCCRQLAAVDD GWCRCGALDHMLSGI YRELGAT-------- EAGHPMAE--

BDAI-1     --RDCCQEVANISN EWCRCGDLGSMLRSV YAALGVG-------- GGPEE-----

WMAI-1     --RDCCQQLADINN EWCRCGDS-SMLRSV YQELGVR-------- EGKE------

WDAI-1     --RDCCQQLADIS- EWPRCGALYSMLDSM YKEHGVS-------- EGQAGTG---
WDAI-2     --RDCCQQLAHIS- EWCRCGALYSMLDSM YKEHGAQ-------- EGQAGTG---
WDAI-3     --RD-CQQ------ ---------------- ---------------- ----------

WTAI-CM1   PRDRCCKELYDAS- QHCRCEAVRYFIGR- -RSDPN--------- SGVLK-----
WTAI-CM2   PRDRCCKELYDAS- QHCRCEAVRYFIGR- -TSDPN--------- SGVLK-----
WTA1-CMa   PKDRCCQELDEAP- QHCRCEAVRYFIGR- -RSHPD--------- WSVLK-----

WTAI-CM16  AKQQCCGELANIP- QQCRCQALRYFMGP- -KSRPD--------Q SGLM------
BTAI-CMb   AKQQCCGELANIP- QQCRCQALRFFMGR- -KSRPD--------Q SGLM------
WTAI-CM17  AKQECCEQLANIP- QQCRCQALRYFMGP- -KSRPD--------Q SGLM------

WTAI-CM3   AKLYCCQELAEIS- QQCRCEALRYFIALP VPSQPVDPRSGNVGE SGLI------
BTAI-CMd   AKLYCCQELAKIP- QQCRCEALRYFMALP VPSQPVDPSTGNVGQ SGLM------
```

C

```
BTI-CME    DSPNCPRERQTSYAA NLVTPQECNLGTIHG S-----AYCPELQPG YGVVL
RTI        DMPNCPRVTQRSYAA TLVAPQECNLPTIHG S-----PYCPTLQAG Y
RBI        DLPGCPRQVQRAFAP KLVTEVECNLATIHG G-----PFCLSLLGA GE
MTI        DLPGCPREVQRGFAA TLVTEAECNLATISG V-----AECPWILGG GTMPSK
PUP-23     DRRDCPREEQPAFAA TLVTAAECNLSSVQE P-----GVRLVLLAD G
```

```
PUP-88      DRRDCPREAQREFAA TLVTAAECNLPTVS- ------GVGSTLGAT GRWMTIELPK

BTI-CMc     AT-ECPAEVKRDFAR TLALPGQCNLPAIHG G-----AYCVFP

BMAI-1      VFPGCRTEV--MDRA VASLPAVCNQYIPNT NGT--DGVCY--WLS YYQPPRQMSSR
RAP         VFPGCRRGD--LERA AASLPAFCNVDIPNG PGG---VVCY--WLG YPRTPRTGH

BDAI-1      VFPGCQKDV--MKLL VAGVPALCNVPIPNE A-AGTRGVCY--WSA STDT

WMAI-1      VLPGCRKEV--MKLT AASVPEVCKVPIPNP SGD-RAGVCYGDWAA YPDV

WDAI-1      AFPSCRREV--VKLT AASITAVCRLPIVVD ASGDGAYVCK-DVAA YPDA
WDAI-2      AFPRCRREV--VKLT AASITAVCRLPIVVD ASGDGAYVCK-DVAA YPDA

WTAI-CM1    DLPGCPREPQRDFAK VLVTSGHCNVMTVHN A-----PYCLGLDI
WTAI-CM2    DLPGCPREPQRDFAK VLVTPGHCNVMTVHN T-----PYCLGLDI
BTAI-CMa    DLPGCPKEPQRDFAK VLVTPGQCNVLTVHN A-----PYCLGLDI

WTAI-CM16   ELPGCPREVQMDFVR ILVTPGYCNLTTVHN T-----PYCLAMEES QWS
WTAI-CM17   ELPGCPREVQMNFVP ILVTPGYCNLTTVHN T-----PYCLGMEES QWS
BTAI-CMb    ELPGCPREVQMDFVR ILVTPGFCNLTTVHN T-----PYCLAMDEW QWNRQFCSS

WTAI-CM3    DLPGCPREMQWDFVR LLVAPGQCNLATIHN V-----RYCPAVEQP LWI
BTAI-CMd    DLPGCPREMQRAFVR LLVAPGQCNLATIHN V-----RYCPAVEQP LWI
```

A barley tetrameric α-amylase inhibitor has been characterized and its subunits have been identified as the previously described proteins CMa, CMb and CMd (Sanchez-Monge *et al.*, 1986). The CMd subunit, the most hydrophobic of the three, is present in two copies in the tetramer. The complete sequences of WTAI-CMa, WTAI-CMb and WTAI-CMd have been deduced from the corresponding cDNAs (Halford *et al.*, 1988; Medina *et al.*, 1993; Paz-Ares, *et al.*, 1986; Rasmussen and Johansson, 1992). Tetrameric inhibitors have also been characterized in hexaploid wheat *Triticum aestivum* (genomes AABBDD), as well as in tetraploid *Triticum turgidum* (AABB) and in diploid *Triticum tauschii* (DD) and cDNAs corresponding to the three types of subunits have been cloned (García-Maroto *et al.*, 1990).

2. GENETICS.

2.1 Chromosomal locations of inhibitor genes.

The trypsin/α-amylase inhibitors are encoded by a multi-gene family dispersed over several chromosomes both in wheat and in barley. Present knowledge of the genomic organization of this family is summarized in Table 1. Five out of the seven homoeologous chromosome groups carry genes coding for these inhibitors. Genes in the A genome of hexaploid wheats appear to be silenced (pseudogenes) (García-Maroto *et al.*, 1990).

2.2 Gene expression

In all the cases where a complete cDNA has been characterized, the mature protein is preceded by a typical signal peptide of aproximately 30 amino acids, which is in agreement with the observation that the synthesis of these inhibitors takes place in membrane-bound polysomes as pre-proteins that are co-translationally processed (Paz-Ares *et al.*, 1983).

The effects of high-lysine mutations on the expression of different genes from this family in barley have been investigated. The most remarkable effect concerns the gene for trypsin inhibitor BTI-CMe, which is regulated in *trans* by the *Lys3a* locus. The CMe protein is present in the mature endosperm of the mutant Risϕ 1508 at less than 2-3% of the wild type level, and the steady state level of the CMe-mRNA is about 1% (Lazaro *et al.*, 1985; Rodriguez-Palenzuela *et al.*, 1989; Royo *et al.*, 1996; Salcedo *et al.*, 1984). Southern blot analysis of wheat-barley addition lines has shown that chromosome 3H of barley carries the gene for CMe. One or two copies of the CMe gene per haploid genome have been estimated both in the wild type and in the mutant and DNA restriction patterns are identical in both stocks, so neither a change in copy number nor a major rearrangement of the structural gene account for the markedly decreased expression. The mutation at the *Lys3a* locus in Risϕ 1508 has been previously mapped in chromosome 5H. A single dose of the wild type allele at this locus (*Lys3a*) restores the expression of gene CMe in chromosome 3H to normal levels (Rodriguez-Palenzuela *et al.*, 1989).

The gene *Itr1*, encoding trypsin inhibitor BTI-CMe from barley, has been functionally analysed by transient expression in protoplasts derived from different barley tissues showing that, under these conditions, its promoter retains both its endosperm specificity and its regulation in *trans* by the *Lys3a* gene (Royo *et al.*, 1996). The proximal promoter extending 343 bp upstream of the initiation codon is sufficient to confer endosperm-specific expression in wild-type protoplasts, whereas expression in protoplasts from the Risϕ1508 *lys3a* mutant was less than 5 % of that in wild type protoplasts (barley cv.

Bomi). Nuclear proteins extracted from the two types of endosperm gave differential patterns in gel retardation experiments. Several transcription factors belonging to the bZIP class have been isolated from barley and their possible involvement in the regulation of the *Itr1* gene is currently being investigated (Vicente-Carbajosa *et al.*, 1998; Oñate *et al.*, in preparation).

Table 1 Chromosomal Location of Genes Encoding Trypsin/α-Amylase Inhibitors

Inhibitory activity against	Aggregation	Protein	Gene	Chromosome, genome, arm(*)
TRYPSIN	monomeric	BTI-CMe	*Itr1*	3HS
		RTI	*Itr-R1*	3R
		BTI-CMc	*Itr2*	7HS
α -AMYLASE	monomeric	BMAI-1	*Iam1*	2H
		WMAI-1 (syn. 0.28)	*Imha-D1*	6DS
		WMAI-2	*Imha-B1*	6BS
α -AMYLASE	homodimeric	BDAI-1	*Iad1*	6H
		WDAI-1 (syn. 0.53)	*IdhaB1.1*	3BS
		WDAI-2 (syn. 0.19)	*IdhaD1.1*	3DS
		WDAI-3	*IdhaB1.2*	3BS
α -AMYLASE	tetrameric			
1st SUBUNIT		BTAI-CMa	*Iat1*	7HS
		WTAI-CM1	*IthaD1*	7DS
		WTAI-CM2	*IthaB1*	7BS
2nd SUBUNIT		BTAI-CMb	*Iat2*	4HL
		WTAI-CM16	*IthaB2*	4BS
		WTAI-CM17	*IthaD2*	4DS
3rd SUBUNIT		BTAI-CMd	*Iat3*	4HL
(2 copies)		WTAI-CM3B	*IthaB3*	4BS
		WTAI-CM3D	*IthD3*	4D
UNKNOWN				
clone pUP23	-	-	-	6HL
clone pUP88	-	-	-	6AL
				6BL
				6DL

(*)S= short, L= long chromosome arms; B,D= wheat (genomes AABBDD); H= barley (genome HH); R= rye (genome RR).

3. BIOLOGICAL PROPERTIES.

3.1 Inhibition of insect enzymes

Barley BTI-CMe and its maize homologue (MTI) are not only active against trypsin but also against Hageman factor XII-a of the blood-clotting cascade, and BTI-CMe is also active against Kallikrein (Chong and Reeck, 1987). BTI-CMe is inactive against chymotrypsin, papain, pepsin, bacterial and fungal proteases and the endogenous barley proteases (Mikola and Soulinna, 1969), as well as against α-amylases (Barber *et al.*, 1986b), which is in contrast with the bifunctional activity of its homologues from maize and ragi (Chen *et al.*, 1992; Shivaraj and Pattabiraman, 1981).

The monomeric, dimeric and tetrameric α-amylase inhibitors from wheat and barley differentially inhibit α-amylases from different origins. In general, the amylase inhibitors purified from barley are less effective than those from wheat. The wheat dimeric class is more active towards the α-amylase from human saliva than against those from insect pests (see Carbonero *et al.*, 1993). The enzymes from both *Tenebrio molitor* (Coleoptera) and *Ephestia kuehniella* (Lepidoptera) are significantly more sensitive than human salivary α-amylase to the monomeric wheat inhibitor WMAI-1. A given inhibitor class may also discriminate whithin an insect group: the *Leptinotarsa decemlineata* enzyme is more affected by the homodimeric inhibitors than by the monomeric ones, while the opposite is true for the α-amylase of *Tenebrio molitor*, both coleopterous insects (Gutierrez *et al.*, 1990). The Lepidoptera seem to be more susceptible to the tetrameric inhibitors than to the monomeric or dimeric ones (Gutierrez *et al.*, 1993).

Although no *in vivo* function is known for the cereal trypsin/α-amylase inhibitor family, the following aspects can be noted:

i) Different inhibitors show different specificities towards enzymes from different insects (Gutierrez *et al.*, 1990; 1993).

ii) Considerable intra- and interspecific variation in inhibitor levels has been observed (Gomez *et al.*, 1989; Kirsi and Ahokas, 1983).

iii) Gene silencing has occurred in some cases (Garcia-Maroto *et al.*, 1990).

These characteristics, together with the low genetic variability at given loci, suggest that these proteins are involved in plant defense, probably as components of the non-host resistance mechanism rather than in relation to more specific interactions. More direct evidence for a defense role stems from experiments with insect pests. Thus, insects which are able to feed on wheat endosperm have unusually high levels of α-amylase (Gutierrez *et al.*, 1990). High inhibitor concentrations in an artifitial diet were required to affect the development of larvae of *Tribolium confusum*, a storage pest of wheat

products, while quite low concentrations were effective against *Callosobruchus maculatus*, a pest of legume seeds (Gatehouse *et al.*, 1986). More recently, transgenic tobacco plants expressing the inhibitors BTI-CMe from barley or WMAI-1 (syn. 0.28) from wheat, under the control of the 35S promoter, have been found to be lethal in leaf-disc assays to larvae from Lepidoptera, such as *Agrotis ipsilon* and *Spodoptera littoralis* (Table2).

Table 2 Insect feeding assays in transgenic (Ro) tobacco leaves

Transgenic plant expressing	Plant no	% Mortality (L_1 to L_3)	
		Agrotis ipsilon	*Spodoptera littoralis*
BTI-CMe	2	22.6	-
	3	62.6*	37.1
	4	52.5*	32.8
	6	30.0	-
WMAI-1	6	72.7*	23.0
	8	17.4	-
	9	38.3	-
	12	65.7*	39.2
Untransformed tobacco	-	15.0	10.5
Artificial diet	-	14.5	12.0

3.2 Allergenic properties

The allergic asthma of workers with occupational exposure to cereal flour is due in part to the allergenic properties of some members of this protein family. Most of these inhibitors isolated from wheat, barley and rice are recognized by specific IgE when tested with sera from allergic patients (Barber *et al.*, 1989; García-Casado *et al.*, 1995; 1996; Gomez *et al.*, 1990; Izumi *et al.*, 1992; Sanchez-Monge *et al.*, 1992; 1996a,b). However, their IgE-binding capacities *in vitro* are very different (Sanchez-Monge *et al.*, 1992). In wheat and barley the glycosylated forms of WTAI-CM16, BTAI-CMb and BMAI-1 have been found to be the most prominent allergens, both *in vitro* and *in vivo* (García-Casado *et al.*, 1995; Sanchez-Monge *et al.*, 1992). cDNA clones encoding these three proteins have also been isolated (García-Maroto *et al.*, 1990; Medina *et al.*,1993; Mena *et al.*, 1992).

4. STRUCTURE

4.1 The reactive site of α-amylase inhibitors

Although, as already indicated, the reactive sites of trypsin inhibitors have been known for some time, it had been speculated that the ability to inhibit α-amylases was mediated by the carbohydrate moieties of the glycosylated members of this family. This question has been finally clarified through mutagenesis. Inhibitor WMAI-1 produced in *Escherichia coli* using the PT7-7 expression vector had the same specific activity towards the α-amylase from the insect *Tenebrio molitor* as the native WMAI-1 purified from wheat. This confirms that the native inhibitor, although presenting a putative *N*-glycosylation site (asparagine-proline-serine), is not glycosylated and contradicts the previous claim that a glycosyl moiety was essential for inhibition (Silano *et al.*, 1977). Site-directed mutagenesis of different regions of the inhibitor (Fig. 2) has shown that modifications of the highly conserved *N*-terminal sequence (serine-glycine-proline-tryptophan) increased the pre-incubation time required for maximum activity, while insertions in the middle of the B domain (position 58) led to inactivation (García-Maroto *et al.*, 1991). When the disulphide-bridge structure of this inhibitor was subsequently established (Poerio *et al.*, 1991), position 58 was found to be close to the *N*-terminus (Fig. 2), which suggests that both regions are part of the reactive site.

4.2 Disulphide bridges and 3D structure determination.

The location of the disulphide bridges within these cysteine-rich molecules has been investigated to a limited extent. The four-bridge structure (9 cysteines) of the wheat dimeric amylase inhibitor WDAI-1 (syn. 0.53), represented in Fig.3, was described by Maeda *et al.* (1983). The presence of one additional cysteine in the sequence of the wheat monomeric amylase inhibitor WMAI-1 (syn. 0.28) was found by Poerio *et al.* (1991) to imply not only the formation of a fifth disulphide bridge, but a general rearrangement of the disulphide structure (Fig. 3). More recently, the 3D structure of the RBI bifunctional α-amylase inhibitor from ragi has been reported (Strobl *et al.*, 1995). This inhibitor has the same disulphide structure as WMAI-1 (Fig. 3). The RBI inhibitor consists of a globular four-helix motif with a simple "up-and-down" topology (Strobl *et al.*, 1995) and there is an antiparallel β-sheet motif between the 3rd and the 4th helices (Plate 10). A location of the putative α-amylase binding site on the face of the molecule opposite to the trypsin-binding loop has been postulated (Strobl *et al.*,1995). Barley trypsin inhibitor

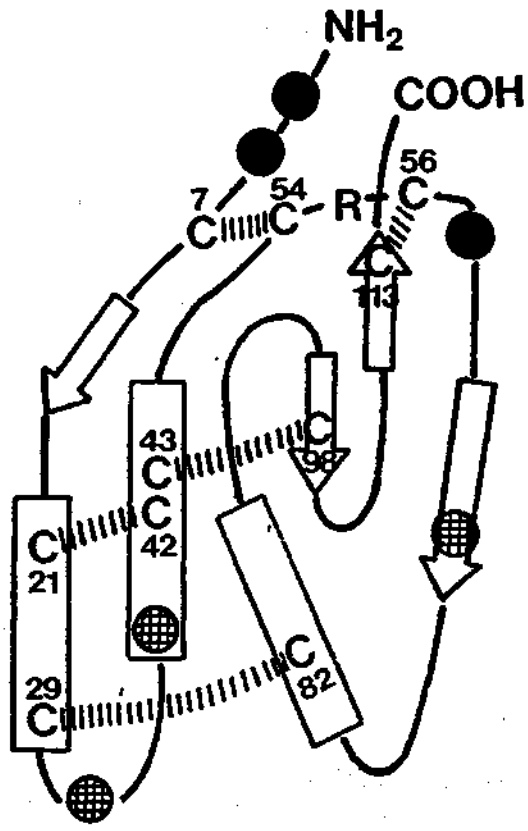

Fig. 2. Schematic view of wheat monomeric α-amylase inhibitor WMAI-1 (syn. 0.28), indicating mutagenised points. Sites where mutagenesis leads to inactivation or diminished affinity towards the α-amylase of *Tenebrio molitor* are depicted as black circles and the mutagenised sites not leading to inactivation as reticulate circles. Beta-sheets (⇒) and alpha-helixes (▯) in the predicted secondary structure are indicated. Disulphide bridges are presented as (llllllll).

BTI-CMe, which has ten cysteines and 55% coincident (69% similar) residues with RBI, is likely to have the same disulphide structure.

5. EVOLUTION

The evolution of this dispersed multigene family raises a number of interesting issues that have been only investigated in a preliminary way:

i) Dispersal of the gene family over several chromosomes of a given genome must have involved both intra-chromosomal duplications and inter-chromosomal translocations. Although some degree of synteny is observed in the gene locations among species, there are significant variations between

closely-related species, such as wheat and barley (Fig. 1; Table 1). A well-studied case of intra-chromosomal duplication is that affecting the *Itr1* locus in *Hordeum vulgare* and *H. spontaneum*. A single active *Itr1* gene is present in most *H. vulgare* cultivars, whereas a duplication of this locus exists in some *H. vulgare* cultivars and in most *H. spontaneum* accessions, and the duplication shows divergent phenotypes within the latter species leading both to active trypsin inhibitors or to pseudogenes (Molina-Cano *et al.*, 1987; Royo *et al.*, submitted; Salcedo *et al.*, 1984;). As to the possible inter-chromosomal dispersal mechanism, it is to be noted that the *Itr1* gene has been shown to be located next to the long terminal repeat of the "copia-like" retro-transposon *Bare-1*, which suggests that transposition may have played a role in the dispersal of the members of this multigene family (Royo *et al.*, 1996).

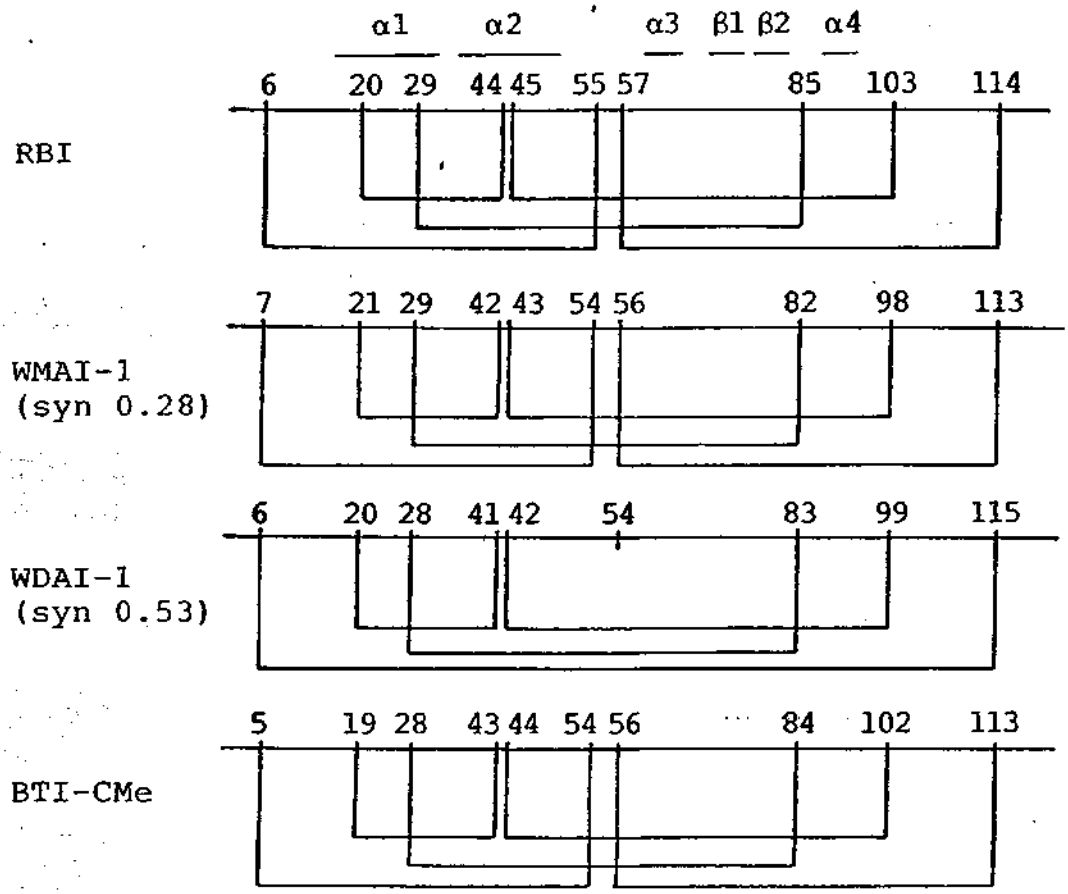

Fig. 3. Schematic representation of the disulphide bond patterns of RBI, WMAI-1 (syn. 0.28) and WDAI-1 (syn. 0.53) and postulated disulphide pattern for barley trypsin inhibitor BTI-CMe.

ii) The possible coevolution of the plant inhibitors and the insect enzymes is suggested indirectly by the available evidence. Little intraspecific variability is observed in the inhibitor pattern, as judged by two-dimensional

electrophoresis. However, there are sharp pattern differences between closely-related species. This is consistent with the idea that each inhibitor pattern represents a specific response to the main pests of the area of distribution of each species. There is also evidence of specificity changes in closely-related inhibitors. Thus, barley inhibitor BTI-CMc, which is a weaker trypsin inhibitor than BTI-CMe, is more closely related in its *N*-terminal sequence (21 out of 29 coincident amino acids) to one of the subunits of the tetrameric inhibitor of α-amylase (BTAI-CMa) than to any of the cereal trypsin inhibitors described so far (Barber *et al.*, 1986a; García-Olmedo *et al.*, 1992; Paz-Ares *et al.*, 1983; Rodriguez-Palenzuela *et al.*, 1989). However the residues deduced from its cDNA at the reactive site place are valine-proline-leucine-leucine, and the rest of the sequence differs sharply from that of BTI-CMe (Figure 1). Also relevant in this context is the above described ability of different α-amylase inhibitors of this family to discriminate among the enzymes from different insect taxa.

iii) The evolution of the aggregative properties of inhibitor subunits also deserves attention. It has been recently reported that two very similar subunits with identical *N*-terminal amino acid sequences (23 residues) have strikingly different properties: one is a subunit of a tetrameric inhibitor that is active against α-amylases from the insect *Tenebrio molitor* (Coleoptera), but not towards α-amylases from other sources, such as *Ephestia kuehniella* (Lepidoptera) or human saliva, whereas the other aggregates as a homodimer and is only active against the human enzyme (García-Casado *et al.*, 1996). A second example is that of the barley dimeric inhibitor BDAI-1, whose chromosomal location and amino acid sequence suggest that this molecule is evolutively closer to the wheat monomeric than to the wheat dimeric inhibitors (Figure 1; Table 1). This would imply that BDAI-1 is a diverged form of the wheat monomeric inhibitors that has acquired the ability to self-associate (Mena *et al.*, 1992).

iv) The existence of hetero-tetrameric inhibitors suggests a molecular model for intergenome heterosis in alloploids. For example, single tetrameric species were observed in *T. tauschii* (subunits CM1, CM3D, CM17) and in *T. turgidum* (CM2, CM3D, CM16), while multiple tetrameric species were observed in *T. aestivum*, resulting from combinations of the subunits contributed by its two parental species (Gomez *et al.*, 1989). The three types of subunits were required for significant activity although binary mixtures involving subunit WTAI-CM1 (or the corresponding barley BTAI-CMa) also had some activity. Additional combinations of subunits were also reconstituted and their inhibitory activities ranged from 144% (CM1, CM3B, CM17) to 33% (CM2, CM3D, CM17) compared to the activity of the reconstituted inhibitor from *T. tauschii*. This, together with the established chromosomal locations of these genes, fit a model of alloploid heterosis at the molecular level (García-Maroto *et al.*, 1990; Gomez *et al.*, 1989).

ACKNOWLEDGEMENTS

This work was financed by grants Bio96-2303 and PB92-0325 from the Comision Interministerial de Ciencia y Tecnología (CICYT) Spain. The technical assistance of L. Lamoneda is acknowledged.

REFERENCES

Adachi, T., Izumi, H., Yamada, T., Tanaka, K., Takeuchi, S., Nakamura, R. and Matsuda, T. (1993) Gene structure and expression of rice seed allergenic proteins belonging to the α-amylase/trypsin inhibitor family. *Plant Molecular Biology* **21**, 239-248.

Barber, D., Sanchez-Monge, R., García-Olmedo, F., Salcedo, G. and Mendez, E. (1986a) Evolutionary implications of sequential homologies among members of the trypsin/α-amylase inhibitor family (CM-proteins) in wheat and barley. *Biochimica et Biophysica Acta*. **873**, 147-151.

Barber, D., Sanchez-Monge, R., Mendez, E., Lazaro, A., García-Olmedo, F. and Salcedo, G. (1986b) New α-amylase and trypsin inhibitors among the CM-proteins of barley (Hordeum vulgare). *Biochimica et Biophysica Acta*. **869**, 115-118.

Barber, D., Sanchez-Monge, R., Gomez, L., Carpizo, J., Armentia, A., Lopez-Otín, C., Juan, F. and Salcedo, G. (1989) A barley flour inhibitor of insect α-amylase is a major allergen associated with baker's asthma disease. *FEBS Letters* **248**, 119-122.

Carbonero, P., Salcedo, G., Sanchez-Monge, R., García-Maroto, F., Royo, J., Gomez, L., Mena, M., Medina, J. and Diaz, I. (1993) A multigene family from cereals which encodes inhibitors of trypsin and heterologous α-amylases. In: "Innovations in Proteases and their Inhitbitors" (F.X. Avilés, ed.) pp. 333-348. Walter de Gruyter, Berlin and New York.

Campos, F.A.P. and Richardson, M. (1983) The complete amino acid sequence of the bifunctional α-amylase/trypsin inhibitor from seeds of ragi (Indian finger millet: Eleusine coracana, Goertn). *FEBS Letters* **152**, 300-304.

Chen, M.S., Feng, G.H., Zen, K.C., Richardson, M., Valdesrodriguez, S., Reeck, G.R. and Kramer, K.J. (1992) α-Amylases from three species of stored grain Coleoptera and their inhibition by wheat and corn proteinaceous inhibitors. Insect *Biochemistry and Molecular Biology* **22**, 261-268.

Chong, G.L. and Reeck, G.R. (1987) Interaction of trypsin, β-factor XIIa and plasma kallikrein with a trypsin inhibitor isolated from barley seeds: a comparison with the corn inhibitor of activated Hageman factor. *Thrombosis Research* **48**, 211-221.

García-Casado, G., Armentia, A., Sanchez-Monge, R., Sanchez, L.M., Lopez-Otín, C. and Salcedo, G. (1995) A major baker's asthma allergen from rye flour is considerably more active than its barley counterpart. *FEBS Letters* **364**, 36-40.

García-Casado, G., Sanchez-Monge, R., Puente, X.S. and Salcedo, G. (1996) Divergence in properties of two closely related α-amylase inhibitors of barley. *Physiologia Plantarum* **98**, 523-528.

García-Maroto, F., Carbonero, P. and García-Olmedo, F. (1991) Site-directed mutagenesis and expression in Escherichia coli of WMAI-1, a wheat monomeric inhibitor of insect α-amylase. *Plant Molecular Biology* **17**, 1005-1011.

García-Maroto, F., Maraña, C., Mena, M., García-Olmedo, F. and Carbonero, P. (1990) Cloning of cDNA and chromosomal location of genes encoding the three types of

subunits of the wheat tetrameric inhibitor of insec α-amylase. *Plant Molecular Biology* **14**, 845-853.

García-Olmedo, F., Salcedo, G., Sánchez-Monge, R., Gomez, L., Royo, J. and Carbonero, P. (1987) Plant proteinaceous inhibitors of proteinases and α-amylases. In: Oxford Surveys of Plant Molecular and Cell Biology. Vol. 4. (B. Miflin, ed.) pp. 275-334. Oxford University Press-ISPMB, New York.

García-Olmedo, F., Salcedo, G., Sanchez-Monge, R., Hernandez-Lucas, C., Carmona, M.J., Lopez-Fando, J.J., Fernandez, J.A., Gomez, L., Royo, J., García-Maroto, F., Castagnaro, A. and Carbonero, P. (1992) Trypsin/α-amylase inhibitors and thionins: possible defence proteins from barley. In: "Barley: Genetics, Biochemistry, Molecular Biology and Biotechnology" (P.R. Shewry, ed.) pp. 335-350. CAB International. Wallingford U.K.

Gatehouse, A.M.R., Fenton, K.A., Jepson, I. and Pavey, D.J. (1986) The effects of α-amylase inhibitors on insect storage pests: inhibition of α-amylase *in vitro* and effects on development in vivo. *Journal of the Science of Food and Agriculture* **37**, 727-734.

Gomez, L., Martin, E., Hernandez, D., Sanchez-Monge, R., Barber, D., Pozo, V., Andres, B., Armentia, A., Lahoz, C., Salcedo, G. and Palomino, P. (1990) Members of the α-amylase inhibitors family from wheat endosperm are major allergens associated with baker's asthma. *FEBS Letters* **261**, 85-88.

Gomez, L., Sanchez-Monge, R., García-Olmedo, F. and Salcedo, G. (1989) Wheat tetrameric inhibitors of insect α-amylase: alloploid heterosis at the molecular level. *Proceedings of the National Academy of Sciences USA.* **86**, 3242-3246.

Gomez, L., Sanchez-Monge, R., Lopez-Otín, C. and Salcedo, G. (1991) Wheat inhibitors of heterologous α-amylases. *Plant Physiology* **96**, 768-774.

Gutierrez, C., Sanchez-Monge, R., Gomez, L., Ruiz-Tapiador, M., Castañera, P. and Salcedo, G. (1990) α-Amylase activities of agricultural insect pests are specifically affected by different inhibitor preparations from wheat and barley endosperm. *Plant Science* **72**, 37-44.

Gutierrez, C., García-Casado, G., Sanchez-Monge, R., Gomez, L., Castañera, P. and Salcedo, G. (1993) Three inhibitor types from wheat endosperm are differentially active against α-amylases of Lepidoptera pests. *Entomology Experimental and Applicata* **66**, 47-52.

Halford, N.G., Morris, N.A., Urwin, P., Williamson, M.S., Kasarda, D.D., Lew, E.J.L., Kreis, M. and Shewry, P. (1988) Molecular cloning of the barley seed protein CMd: a variant member of the α-amylase/trypsin inhibitor family of cereals. *Biochimica et Biophysica Acta* **950**, 435-440.

Izumi, H., Adachi, T., Fujii, N., Matsuda, T., Nakamura, R., Tanaka, K., Urisu, A. and Y. Kurosawa. (1992) Nucleotide sequence of a cDNA clone encoding a major allergenic protein in rice seeds: homology of the deduced aminoacid sequence with members of α-amylase/trypsin inhibitor family. *FEBS Letters* **302**, 213-216.

Kashlan, N. and Richardson, M. (1981) The complete amino acid sequence of a major wheat protein inhibitor of α-amylase. *Phytochemistry* **20**, 1781-1784.

Kirsi, M. and Ahokas, H. (1983) Trypsin inhibitor activities in the wild progenitor of barley. *Phytochemistry* **22**, 2739-2740.

Lazaro, A., Barber, D., Salcedo, G., Mendez, E. and Garcia-Olmedo, F. (1985) Differential effects of high-lysine mutations on the accumulation of individual members of a group of proteins encoded by a disperse multigene family in the endosperm of barley (Hordeum vulgare L.). *European Journal of Biochemistry* **149**, 617-623.

Lazaro, A., Rodriguez-Palenzuela, P., Maraña, C., Carbonero, P. and García-Olmedo, F. (1988a) Signal peptide homology between the sweet protein thaumatin II and unrelated cereal α-amylase/trypsin inhibitors. *FEBS Letters* **1**, 147-150.

Lazaro, A., Sanchez-Monge, R., Salcedo, G., Paz-Ares, J., Carbonero, P. and García-Olmedo, F. (1988b) A dimeric inhibitor of insect α-amylase from barley. *European Journal of Biochemistry* **172**, 129-134.

Lyons, A., Richardson, M., Tatham, A.S. and Shewry, P.R. (1987) Characterization of homologous inhibitors of trypsin and α-amylase from seeds of rye (Secale cereale L.) *Biochimica et Biophysica Acta*. **915**, 305-313.

Maeda, K., Wakabayashi, S. and Matsubara, H. (1983) Disulfide bridges in an α-amylase inhibitor from wheat kernel. *Journal of Biochemistry* **94**, 865-870.

Maeda, K., Wakabayashi, S. and Matsubara, H. (1985) Complete amino acid sequence of an α-amylase inhibitor in wheat kernel (0.19 inhibitor). *Biochimica et Biophysica Acta* **828**, 213-221.

Mahoney, W.C., Hermodson, M.A., Jones, B., Powers, D.D., Corfman, R.S. and Reeck, G.R. (1984) Amino acid sequence and secondary structural analysis of the corn inhibitor of trypsin and activated Hageman factor. *Journal of Biological Chemistry* **259**, 8412-8416.

Medina, J., Hueros, G. and Carbonero, P. (1993) Cloning of cDNA, expression and chromosomal location of genes encoding of three types of subunits of the barley tetrameric inhibitor of insect α-amylase. *Plant Molecular Biology* **23**, 535-542.

Mena, M., Sanchez-Monge, R., Gomez, L., Salcedo, G. and Carbonero, P. (1992) A major barley allergen associated with baker's asthma disease is a glycosylated monomeric inhibitor of insect α-amylase: cDNA cloning and chromosomal location of the gene. *Plant Molecular Biology* **20**, 451-458.

Mikola, J. and Soulinna, E.M. (1969) Purification and properties of a trypsin inhibitor from barley. *European Journal of Biochemistry* **9**, 555-560.

Molina-Cano, J.L., Fra-Mon, P., Salcedo, G., Aragoncillo, C., Roca de Tagores, F. and García-Olmedo, F. (1987) Morocco as a possible domestication center for barley: biochemical and agromorphological evidence. *Theoretical and Applied Genetics* **73**, 531-536.

Odani, S., Koide, T. and Ono, T. (1983) The complete amino acid sequence of barley trypsin inhibitor. *Journal of Biological Chemistry* **258**, 7998-8003.

Paz-Ares, J., Ponz, F., Aragoncillo, C., Hernandez-Lucas, C., Salcedo, G., Carbonero, P. and García-Olmedo, F. (1983) In vivo and *in vitro* synthesis of CM-proteins (A-hordeins) from barley (Hordeum vulgare L.). *Planta*. **157**, 74-80.

Paz-Ares, J., Ponz, F., Rodriguez-Palenzuela, P., Lazaro, A., Hernandez-Lucas, C., Garcia-Olmedo, F. and Carbonero, P. (1986) Characterization of cDNA clones of the family of trypsin/α-amylase inhibitors (CM-proteins) in barley (Hordeum vulgare L.). *Theoretical and Applied Genetics* **71**, 842-846.

Poerio, E., Caporale, C., Carrano, L., Pucci, P. and Buonocore, V. (1991) Assignment of the five disulfide bridges in an α-amylase inhibitor from wheat kernel by fast-atom bombardment mass spectrometry and Edman degradation. *European Journal of Biochemistry* **199**, 595-600.

Rasmussen, S.K. and Johansson, A. (1992) Nucleotide sequence of a cDNA encoding the barley seed protein CMa: an inhibitor of insect α-amylase. *Plant Molecular Biology* **18**, 423-427.

Rodriguez-Palenzuela, P., Royo, J., Gomez, L., Sanchez-Monge, R., Salcedo, G., Molina-Cano, J.L., García-Olmedo, F. and Carbonero, P. (1989) The gene for trypsin inhibitor CMe is regulated in trans by the *Lys3a* locus in the endosperm of barley (Hordeum vulgare L.). *Molecular and General Genetics* **219**, 474-479.

Royo, J., Diaz, I., Rodriguez-Palenzuela, P. and Carbonero, P. (1996) Isolation and promoter characterization of the barley gene Itr1 encoding trypsin inhibitor BTI-CMe: differential activity in wild type and mutant *Lys3a* endosperm. *Plant Molecular Biology* **31**, 1051-1059.

Royo, J., Gaddour, K., Vicente-Carbajosa, J., Diaz, I and Carbonero, P. (1997) Gene duplication and evolution of the Itr1 locus in cultivated and wild Hordeum spp. (submitted).

Salcedo, G., Fra-Mon, P., Molina-Cano, J.L., Aragoncillo, C. and García-Olmedo, F. (1984) Genetics of CM-proteins (A-hordeins) in barley. *Theoretical and Applied Genetics* **68**, 53-59.

Sanchez de la Hoz, P., Castagnaro, A. and Carbonero, P. (1994) Sharp divergence between wheat and barley at loci encoding novel members of the trypsin/α-amylase inhibitors family. *Plant Molecular Biology* **26**, 1231-1236.

Sanchez-Monge, R., García-Casado, G., Barber, D. and Salcedo, G. (1996a) Interaction of allergens from house-dust mite and from cereal flours: Dermatophagoides pteronyssinus α-amylase (Der p 4) and wheat and rye α-amylase inhibitor. *Allergy* **51**, 176-180.

Sanchez-Monge, R., García-Casado, G., Malpica, J.M. and Salcedo, G. (1996b) Inhibitory activities against heterologous α-amylases and *in vitro* allergenic reactivity of Einkorn wheats. *Theoretical and Applied Genetics* **93**, 745-750

Sánchez-Monge, R., Gomez, L., García-Olmedo, F. and Salcedo, G. (1986). A tetrameric inhibitor of insect α-amylase from barley. *FEBS Letters* **207**, 105-109.

Sanchez-Monge, R., Gomez, L., Barber, D., Lopez-Otín, C., Armentia, A. and Salcedo, G. (1992) Wheat and barley allergens associated with baker's asthma. Glycosylated subunits of the α-amylase-inhibitor family have enhanced IgE-binding capacity. *Biochemical Journal* **281**, 401-405.

Shewry, P.R., Lafiandra, D., Salcedo, G., Aragoncillo, C., García-Olmedo, F., Lew, E.J.-L., Dietler, M.D. and Kasarda, D.D. (1984) *N*-terminal amino acid sequences of chloroform-methanol-soluble proteins and albumins from the endosperm of wheat, barley, and related species. *FEBS Letters* **175**, 359-363.

Shivaraj, B. and Pattabiraman, T.N. (1981) Natural plant enzyme inhibitors. Characterization of an unusual α-amylase/trypsin inhibitor from ragi (Eleusine coracana Geartn). *Biochemical Journal* **193**, 29-36.

Silano, V., Poerio, E. and Buonocore, V. (1977) A model for the interaction of wheat monomeric and dimeric protein inhibitors with α-amylase. *Molecular and Cell Biochemistry* **18**, 87-91.

Strobl, S., Mühlhahn, P., Bernstein, R., Witscheck, R., Maskos, K., Wunderlich, M., Huber, R., Glockshuber, R. and Holak, T.A. (1995) Determination of the three-dimensional structure of the bifunctional α-amylase/trypsin inhibitor from ragi seeds by NMR spectroscopy. *Biochemistry* **34**, 8281-8293.

Vicente-Carbajosa, J., Oñate, L., Lara, P., Diaz, I. and Carbonero, P. (1998) Barley BLZ1: a bZIP transcriptional activator that interacts with endosperm-specific gene promoters. *The Plant Journal* **13**, 629-640.

Wen, L., Huang, J-K., Zen, K.C., Johnson, B.H., Muthukrishnan, S., Mackay, V., Manney, T.R., Manney, M. and Reeck, G.R. (1992) Nucleotide sequence of a cDNA clone that encodes the maize inhibitor of trypsin and activated Hageman factor. *Plant Molecular Biology* **18**, 813-814.