

Extreme Divergence of a Novel Wheat Thionin Generated by a Mutational Burst Specifically Affecting the Mature Protein Domain of the Precursor

Atilio Castagnaro, Carmen Marañón, Pilar Carbonero and Francisco García-Olmedo

*Area de Bioquímica y Biología Molecular. E.T.S
Ingenieros Agrónomos, UPM. E-28040 Madrid, Spain*

A new type of neutral thionin (type V), specifically expressed in developing wheat endosperm, has been found to be encoded by a set of single-copy genes located in the long arms of chromosomes 1A, 1B and 1D, within less than 10,000 base-pairs of those corresponding to the highly basic type-I thionins. Divergence between types I and V has occurred through a process of accelerated evolution that has affected the amino acid sequence of the mature thionin but not the precursor domains corresponding to the N-terminal signal peptide and the long C-terminal acidic peptide. This process involved a deletion and a non-synonymous nucleotide substitution rate equal to the synonymous rate in the thionin sequence.

Keywords: accelerated evolution; antimicrobial peptides; thionins; wheat; gene duplication

1. Introduction

The thionins are a homologous family of cysteine-rich polypeptides of $M_r \approx 5000$ that have been isolated from a wide range of plant taxa and are active against plant pathogens (for reviews, see García-Olmedo *et al.*, 1989, 1990, 1991). Based on the multiple-alignment criteria of Feng & Doolittle (1987), all the available thionin amino acid sequences were classified into four well-defined structural types (I to IV, García-Olmedo *et al.*, 1989, 1990), two of which were present in the cereals wheat and barley. Type I corresponds to the original purothionins from wheat flour (Balls *et al.*, 1942a,b), which were later characterized in barley endosperm (Redman & Fisher, 1969), and type II is represented by thionins first isolated from the leaves of the parasitic plant *Pyrularia pubera* (Vernon *et al.*, 1985) and then identified in those of barley (Bohlman & Apel, 1987; Gausing, 1987). Types III and IV respectively include the viscotoxins from mistletoes and the crambins from the Abyssinian cabbage (see García-Olmedo *et al.*, 1990).

Type-I thionins are characterized by four disulfide bridges and are highly basic. They are synthesized during the first stages of endosperm development as much larger precursors that

undergo at least two processing steps: the cotranslational elimination of a signal peptide and the post-translational excision of a C-terminal acidic peptide (Ponz *et al.*, 1983, 1986; Hernández-Lucas *et al.*, 1986). Type-I thionins are encoded by single-copy genes that are located in the long arms of group-1 chromosomes (Fernández de Caleyá *et al.*, 1976), and have two introns that interrupt the sequence encoding the acidic peptide (Rodríguez-Palenzuela *et al.*, 1988).

Type-II thionins have also four disulfide bridges at the same positions as those of type I, but are less basic, with some negatively charged residues, and their central disulfide loop is one or two residues longer. The biosynthesis of these thionins has not yet been investigated, but the structures of their precursors, which have been deduced from the nucleotide sequences of their cDNAs (Bohlman & Apel, 1987; Gausing, 1987), are identical with that previously described for type-I thionins (Hernández-Lucas *et al.*, 1986; Ponz *et al.*, 1986). Type-II genes have been located in chromosome 6 of barley (Bohlman *et al.*, 1988) and their number per haploid genome has been variously estimated to 9 to 11 (Gausing, 1987) and at 50 to 100 copies (Bohlman *et al.*, 1988). Although genomic clones of type-II thionins have not been described in detail, it seems that they also have two introns (Bohlman *et al.*, 1988). These genes are expressed in young etiolated barley leaves, but not in illuminated ones (Bohlman

& Apel, 1987; Gausing, 1987; Reiman-Philipp *et al.*, 1989). The inhibitory effect of light could be overcome by stress- and pathogen-induced signals (Bohlman *et al.*, 1988; Ebrahim-Nesbat *et al.*, 1989; Fisher *et al.*, 1989).

We now report a neutral thionin type from wheat, designated type V, which has diverged extensively from type I by a process of accelerated evolution specifically affecting the mature protein domain of the precursor. Type-V genes have been located close to type-I genes in the long arms of group-1 chromosomes in the three genomes (ABD) of allohexaploid wheat, *Triticum aestivum* L., and have been found to be expressed in endosperm at the same time as type-I genes.

2. Materials and Methods

(a) Biological material

Hexaploid wheat, *Triticum aestivum* L. (genomes ABD) cultivars Almatense and Chinese Spring, nulli-tetrasomic and ditelosomic lines of cultivar (cv.†) Chinese Spring (gift of E. R. Sears, Columbia, MO, U.S.A.), tetraploid wheat, *T. turgidum* (AB) AP-1, and diploid wheat, *T. monococcum* (A), were used in this study.

(b) cDNA cloning and library screening

The construction and screening of a wheat cDNA library was carried out as described previously (Maraña *et al.*, 1988). Total polysomal RNA from developing endosperm (20 days after pollination) of cv. Chinese Spring was the source of the poly(A)⁺ RNA used. The library was screened using the insert in barley thionin cDNA clone pTH-1 as a probe (Ponz *et al.*, 1986), which was ³²P-labeled by nick translation. Non-stringent conditions (58°C) were used in the screening.

(c) Nucleotide sequencing

The complete sequence of the selected clones was determined using the dideoxy chain termination method (Sanger *et al.*, 1977). Inserts were subcloned into M13 (mp18 and mp19) and single-stranded DNA annealed to the primer by heating at 65°C for 2 min and cooling slowly to room temperature. Sequencing reactions used [α -³⁵S]dATP (>1000 Ci/mmol) as label and the products were separated on 8% polyacrylamide/9 M-urea "wedge" gels. Direct sequencing of DNA amplified by the polymerase chain reaction (PCR) was carried out by a published procedure involving detergent treatment (Bachman *et al.*, 1990). The sequences were analyzed with the Beckman Microgenie software.

(d) DNA blots, RNA blots and hybridization

Genomic DNA was isolated from 7-day-old dark-grown wheat seedlings essentially as described by Murray & Thompson (1980). Restriction digestion, agarose gel electrophoresis and Southern blotting to nylon membranes (Hybond N, Amersham) were performed using standard procedures (Sambrook *et al.*, 1989) and according to the manufacturer's instructions.

Hybridization to nick-translated inserts of appropriate cDNA clones was in 5 × SSPE (0.9 M-NaCl, 0.05 M-NaH₂PO₄ (pH 7.4), 0.005 M-EDTA), 2 × Denhardt's solution (0.04% polyvinylpyrrolidone, 0.04% bovine serum albumin, 0.04% Ficoll), 0.2% SDS, 100 µg salmon sperm DNA/ml, at 65°C or at 58°C, according to the stringency required in each case. The RNA for Northern blots was isolated from developing endosperms, coleoptiles and roots from 7-day-old etiolated or illuminated plants and from mature leaves of illuminated 4-week-old plants. Glyoxal-denatured RNAs (Thomas, 1983) were fractionated by electrophoresis in 1.2% agarose gels, transferred to nitrocellulose filters (Hybond C, Amersham) and hybridized as previously described (Ponz *et al.*, 1986). Probes used in Southern- and Northern-blot experiments were obtained by PCR amplification with primers *a* and *b* (Fig. 4) of the insert in plasmid pTTH20 (probe V) and by electrophoretic purification of the insert in clone pTTH1 encoding α 1-thionin from wheat endosperm (probe I; A. Castagnaro and C. Maraña, unpublished results).

(e) Polymerase chain reaction

DNA amplification was carried out using the Perkin-Elmer Cetus reagent kit and thermal cycler (1 min, 95°C; 1 min, 63°C; 1.5 min, 72°C).

(f) Mutability index

The index of mutability, *Im*, derived by Graur (1985) was calculated using the formula $Im = 0.841 - 5.096 f_{Gly} + 24.145 f_{Asn} - 26.807 f_{Tyr} - 7.398 f_{Val} + 18.219 f_{Phe} - 8.263 f_{Asp} + 7.960 f_{Ile}$.

3. Results

(a) Identification of cDNA clone pTTH20 encoding the precursor of a new type of neutral thionin

A bank of cDNA clones obtained from developing wheat endosperm was screened under low-stringency conditions, using a barley α -thionin cDNA as probe. Nucleotide sequences of all but one of eleven selected clones corresponded to previously described type-I thionins from wheat endosperm (not shown). The eleventh clone, designated pTTH20, encoded a complete thionin precursor (Fig. 1) that showed extreme divergence in its mature protein domain when aligned with known thionin types and therefore represented a new thionin class, designated type V (Fig. 2). This type is quite distant from the other four, but its six cysteine residues are at positions that are conserved with respect to types I and II, while only four of these are found at the same positions in types III and IV (Fig. 2). A relevant feature of the new thionin is a nine amino acid deletion precisely affecting the C terminus of the mature protein domain (Fig. 2). The C-terminal, acidic peptide sequence of type-V thionin is closer to that of type I than to that of type II (65% versus 38% coincident residues).

(b) Type-V genes are located near type-I genes

Southern-blot analysis of DNAs from hexaploid and tetraploid wheat, independently digested with

† Abbreviations used: cv., cultivar; PCR, polymerase chain reaction; kb, 10³ base-pairs; bp, base-pair(s).

SIGNAL PEPTIDE

CCCATTAACTTGTGCCTCTCTGCAAGAGAGCCCGGCAACCATAGCCATGGAGGAGGCGCCAAAGGCTCTGGAGAGTGGGATCGTGTGCCCTACTGGTGGGGCTGGCTGG 115
M G G G Q K G L E S A I V C L L V L G L V L 22

THIONIN

AGCAGGTCCAAGTAGAGGGGTGGACTCCGGCCGAACCCCTTCAAAGTGGGCTCTCAACTCTTGCCTTCTCGCTTAGTACAGTGTCCAGTGGCCAGCTCTGTCTCTG 230
E Q V Q V E G V D C G A N P F K V A C F M S C L L G P S T V F Q C A D F C A C 61

ACIDIC PEPTIDE

TCGACTCCCGGGGGCTGGCCCTCTGTACGCGAGTCCGATGAACCAACGCCATTCAGTATTGCAGCTTGGGATGTAGGTCTTCCGTGTGTGACAACATGATCAACAGCTGAC 345
R L P A G L A S V R S S D E P M A I E Y C S L G C R S S V C D N H I N T A D 99

ScaI

AACAGTACTGAAGAGATGAACCTCTATGTGAACCGCTCCGGTGTCTGACAGTTTCTGTAAAGGGGATACTCTTGGCATCCCTTGATGACTAATAATGATATGCTGG 460
N S T E E M K L Y V K R C G V A C D S F C K G D T L L A S L D D * * 131

TCTTAGATTTCAGAGTCAAAGTATGTCTCACCATCGTGTTCATAAATCAGGATCCCATTCAGAGTATCCAAACCAAGTGTGCAGTTTGTATGTCTGTGGTATTTTCTT 575

TCGTGCTCCATAAATGATATCATGTTGGATGTCTATTTGCTGCTTGTACTGCCATGACTGAGATATAATAATATATGTTGTATCTCTTCTATGGATATAAC 688

Figure 1. Nucleotide sequence of clone pTTH20 (EMBL accession number X61670) and deduced amino acid sequence corresponding to the thionin precursor. The signal peptide, mature thionin and acidic peptide domains of the precursor are indicated. Polyadenylation signals are boxed.

three different restriction enzymes, was carried out using the amplified 5' end of the insert in clone pTTH20 (probe V) and the insert in clone pTTH1 (probe I) as probes (Fig. 3(a)). Under conditions in which no cross-hybridization occurred between the two probes, the same hybridization patterns were obtained for *EcoRI*-digested DNAs with both of them (Fig. 3(a)). In the *BglII* patterns, two bands shared by both wheat species (therefore encoded by genomes A and/or B) hybridized equally to probes I and V, while a third band (encoded by the D genome of hexaploid wheat) appeared with a different mobility with each probe (Fig. 3(a)). Finally, the enzyme *BamHI* gave clearly different patterns with the two probes (Fig. 3(a)). These observations indicated that the two types of loci were within less than 10 kb (1 kb = 1000 base-pairs) from each other and that at least one *BamHI* site existed between the mature protein sequences of types I and V. Analysis of DNAs from hexaploid (genomes ABD), tetraploid (AB) and diploid (A) wheats, digested with *BamHI*, showed that each of the three fragments corresponded to one of the three genomes (Fig. 3(b)). Since type-I loci had been previously located in the long arms of chromosomes 1A, 1B and 1D (Fernández de Caleyra *et al.*, 1976), it

was predicted that the three *BamHI* fragments should be associated respectively with each of these three chromosome regions. This was corroborated through analysis of appropriate nulli-tetrasomic lines of wheat cv. Chinese Spring (Fig. 3(c)). Furthermore, ditelosomic lines lacking respectively the short arms of chromosomes 1A, 1B and 1D, had the normal euploid pattern (not shown).

(c) *Positions of introns are conserved in thionin genes of types I and V*

Genomic DNA corresponding to the insert in clone pTTH20 was amplified by PCR, using primers *a* and *c*, which flanked the coding region (Fig. 4(a)). The amplification product was about 400 base-pairs (bp) longer than that obtained with the cloned cDNA as template, using the same primers (Fig. 4(b)). When primer *c* was substituted by primer *b*, just downstream from the C terminus of the mature protein, a PCR product of the expected size was obtained, whether the amplified genomic DNA or the cloned cDNA was used as template (Fig. 4(b)). These results indicated that introns interrupted the sequence coding for the acidic peptide region of the precursor. Primers *a*, *b* and *c*

SIGNAL PEPTIDE

THIONIN

V
I-1
I-2
II-1
II-2
III
IV

ACIDIC PEPTIDE

V
I-1
I-2
II-1
II-2

Figure 2. Alignment of the following sequences: V, precursor of type-V thionin; I-1, precursor of α -thionin from wheat endosperm (Castagnaro, A. & Marañón, C., unpublished results); I-2, precursor of α -thionin from barley endosperm (Ponz *et al.*, 1986); II-1, II-2, precursors of DB4 and DC4 thionins from barley leaves (Bohman & Apel, 1987); III, mature viscotoxin V-1PS (Samuelsson, 1974); IV, mature crambin (Teeter *et al.*, 1981). Only residues that are identical with the type-V sequence have been boxed.

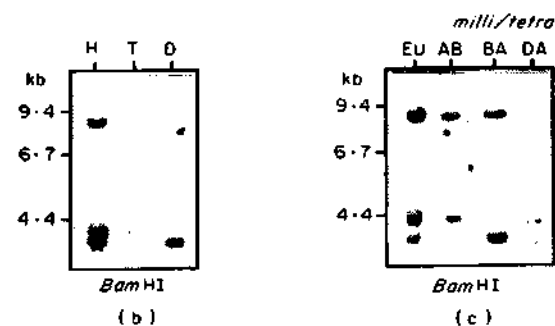
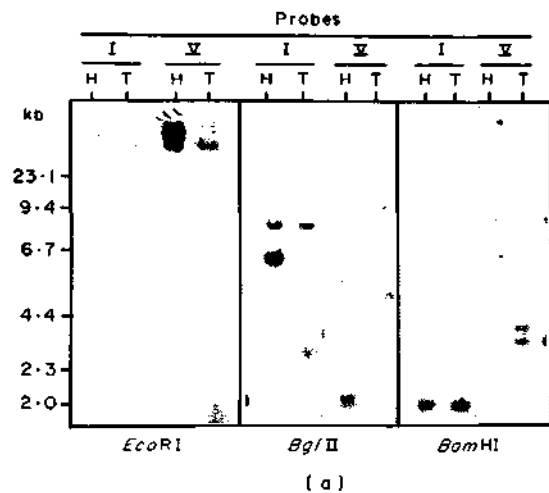


Figure 3. Southern-blot analysis of thionin genes. (a) DNA from hexaploid wheat cv. Almatense (H) and tetraploid wheat API (T) digested with the indicated restriction enzymes and hybridized with a type-I wheat α 1-thionin cDNA (I) and with the signal peptide-mature protein part of the cDNA in clone pTTH20 (V). (b) DNAs from hexaploid (H), tetraploid (T) and diploid (D) wheats hybridized with probe V. (c) DNAs of the following aneuploid stocks from hexaploid cv. Chinese Spring: euploid (EU); null1A-tetra1B (AB); null1B-tetra1A (BA); null1D-tetra1A (DA) hybridized with probe V (nulli means absence of indicated chromosome).

were used for direct sequencing of the amplified products and the *SalI*-*ScaI* fragment (Fig. 4(a)) was subcloned in M13 (mp18 and mp19) and sequenced from both ends. The coding regions of the amplified product were found to be identical with those in clone pTTH20, including the deletion in the C terminus of the mature thionin, and two introns were found at exactly the same positions as in the previously published type-I gene, *Htha* (Rodríguez-Palenzuela *et al.*, 1988). The sequences of the two introns of the new gene, designated *ThV*, have been aligned with those of the *Htha* gene in Figure 5.

In contrast with the good yield obtained in the PCR amplification under stringent conditions when DNA from hexaploid wheat was used as template, very low yields resulted when DNAs from tetraploid or diploid wheats were used with primers *a* and *c*, indicating that clone pTTH20 probably corresponded to the type-V gene located in the D genome of hexaploid wheat.

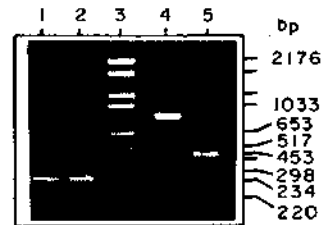
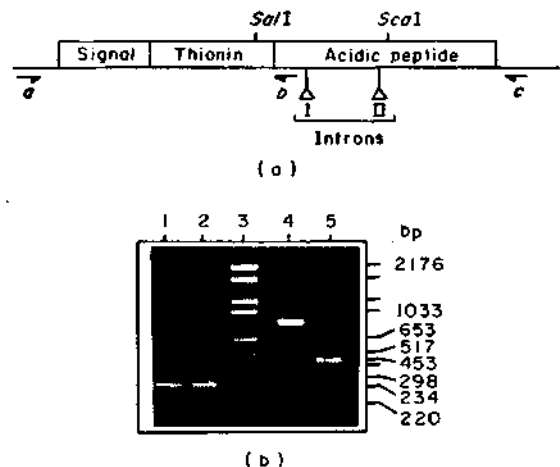


Figure 4. Amplification by PCR of the *ThV* gene. (a) Schematic representation of thionin DNA indicating position and polarity of primers. (b) Electrophoretic separation of the following amplification products: lane 1, amplified DNA in lane 4 reamplified with primers *a* and *b*; lane 2, insert in clone pTTH20 amplified with primers *a* and *b*; lane 3, DNA molecular-weight markers VI (Boehringer-Mannheim); lane 4 DNA from hexaploid wheat cv. Chinese Spring amplified with primers *a* and *c*; lane 5, insert in clone pTTH20 amplified with primers *a* and *c*. Primer *a*: ACTTGTGCCTTCTGCAAGAG (positions 11 to 30 in Fig. 1). Primer *b*: GAGCTGCGTACAGAGGCCAG (positions 266 to 247). Primer *c*: ATCTAAGACCAAGCATATGC (positions 468 to 449).

(d) Endosperm-specific expression of type-V genes

Total RNA was extracted from developing endosperm at different times after pollination, as well as from coleoptiles, and roots from light-grown and etiolated young wheat plants and from mature leaves from light-grown plants. Northern-blot analysis of these RNAs using probes I and V are shown in Figure 6. These results indicated that type-V genes were expressed synchronously with type-I genes in developing endosperm and were not expressed in the other tissues tested, including etiolated coleoptiles in which high expression levels of type-II genes are known to occur (Bohlman & Apel, 1987; Gausing, 1987).

4. Discussion

The thionin encoded in clone pTTH20 significantly differs from previously known types, not only in its amino acid sequence but also in its disulfide structure: cysteine residues 2 and 8, present in types I and II, are absent from this type, through point mutation and deletion, respectively (Fig. 7(a)). Thionins of types I, III and IV have been subjected to extensive structural studies (cited by García-Olmedo *et al.*, 1989), which have shown that all of them have essentially the same three-dimensional shape, both in solution and in crystal form, and that their disulfide bridges are arranged as represented schematically in Figure 7(a). Loss of cysteine residues 2 and 8 in type V disrupts the first

INTRON I

269 375
TthV GTAAAGACAGACAACCTCGCTACTTCTAATTTATATACTATCTCCTTTTCAAAGTAGATGTGATCCCTAAAAAGTGTAATGTCARAACTAAAAAAAATCCAAAGTTTAC
 HthA GTAAAGGGAATGAACCTCTTCTATTCATATTTGCTGATTTCTTCTAAAAATAGATGCCATCAT ATAAAGCGT ATGGCGAGTT AGAGTTTCTAAAGTTTTC

482
TthV CTATTAATACAATAAACTATTCTGAATTTGTTAGTTCATATTTGTATGATAAGATTTCACATGAAGGATTTTTCATTATCTAGAAAAAATCTACCAAAAAATTTGTTG
 HthA CTAAATATAAT AACTTTTTTAATTTTTTGTATGATATTTGTATG AGTTTTTCCATGAAGGTTCCATAAATATACTCGAATTTTTTACCAAAAAATTTTTT

588
TthV TGA AACACAGCTAATAAAGAAATGAATGATAGAGGCCATCCCATACACAAAACATCGACATATTTAAACGGAGGAATTTGTTTATCTCAGATCTACTTCTGTTTGGCAG
 HthA TTCGAAAGGCTAATAAAAAAGCAATGACGGAGACGATGCCAATATCAAGACCGATAATTTATCTGAAGAGAGGAATTTGTTTCTTATTTTACCTTTTATTTGACAG
 TCTCAGGATCCGATTCATCAAAAAAGTTAAGAACGAGTTGCTGTGTTACCGAAAAGTAATTTGTAATAACT TACTCGATTA ATTTACATGCTAGATCT

INTRON II

661 758
TthV GTAATAAAAAATCCATCTGAGTATATTTATATTTTCATTTGTACATGCAAACTAATTA AAAAGTTAGTCTTTGACGCATACTATCCATATTTGGCACCAG
 HthA GTAATAAAAAATTCATCTG 6 TATGTGATTT CCTGTAAACTAATTA AAAAGGCTACTTCTTACACACATAATATGCTATTTGGCACCAG
 TATA

Figure 5. Nucleotide sequences of introns I and II from gene *TthV* aligned with those of gene *HthA* from barley (Rodríguez-Palenzuela *et al.*, 1988). Numbers above sequence correspond to the deposited gene sequence (EMBL accession number X61671).

and second disulfide bridges and potentially allows the formation of a new bridge between the unmatched cysteine residues, 1 and 7 (Fig. 7(a)). Although quite divergent from the other thionins, the new thionin is closer to type I than to the other types, especially if the whole precursor molecules are considered. Furthermore, genes of types I and V are near each other in group-1 chromosomes and have conserved the same tissue specificity and temporal expression, indicating a common origin by duplication of an ancestral gene. This duplication must have occurred before the evolutionary branching out of the diploid species that generated hexaploid wheat. The extreme sequence divergence

between the duplicated loci has led to a highly basic thionin, with ten basic residues out of 45 (type I), and to a neutral one, with two basic and two acidic residues out of 36 (type V). A comparison of genes *TthV* and *HthA* shows that this divergence implies a much higher degree of change affecting the mature protein sequence in comparison with contiguous sequences, including those of introns and of other non-coding regions (Fig. 7(b)). The rates of synonymous (*kS*) and non-synonymous nucleotide substitutions (*kA*) have been estimated for all possible binary combinations between available sequences for the signal peptide, the mature thionin and the acidic peptide of types I, II and V, using a published computer programme (Li *et al.*, 1984). In Figure 7(c), the average *kA/kS* ratios for binary combinations involving type V are represented for all three domains alongside those corresponding to combinations that do not involve type V. The *kA/kS* ratios for the first group of combinations are higher than for the second group; strikingly so in the case of the mature protein domain, where the non-synonymous rate equals the synonymous one. This is a rather singular situation, as it is now well known that synonymous nucleotide substitutions, which do not cause amino acid changes, occur between two and twenty times more rapidly than non-synonymous substitutions (Li *et al.*, 1984; Graur, 1985; Graur & Li, 1988). These differences in substitution rates have been ascribed to selective constraints related to structure-function relationships. Only a limited number of cases have been reported that are exceptions to this rule, all of them involving families of proteinase inhibitors (Fioretti *et al.*, 1985; Creighton & Charles, 1987; Hill & Hastie, 1987; Laskowski *et al.*, 1987a,b; Graur & Li, 1988). Positive Darwinian selection has been invoked to explain accelerated changes taking place at functionally relevant sequences (Hill & Hastie, 1987), but an alternative explanation consistent

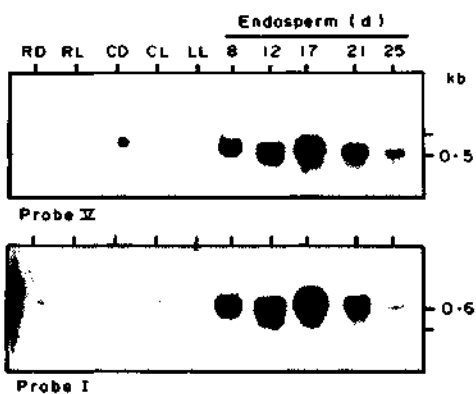


Figure 6. Northern-blot analysis of total RNAs extracted from the indicated organs of hexaploid wheat cv. Chinese Spring. RD, roots from dark-grown 7-day-old seedlings; RL, roots from light-grown 7-day-old seedlings; CD and CL, coleoptiles from dark- and light-grown 7-day-old seedlings; LL, leaves from light-grown 4-weeks plants; 8 to 25, endosperm collected at the indicated days after pollination. Probes I and V as in Fig. 3. A ribosomal DNA probe was used to ascertain that there was no substantial variation (<30%) in the amounts of total RNA applied in each lane.

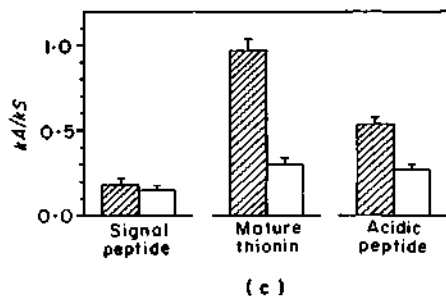
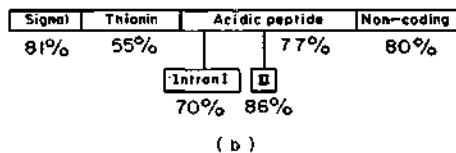
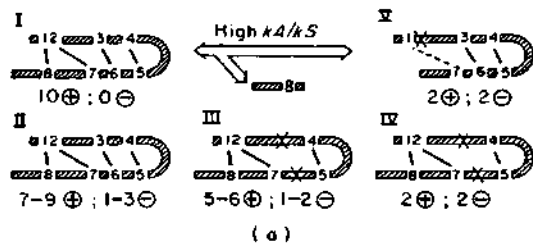


Figure 7. Evolutionary aspects of type-V thionin. (a) Schematic representation of the disulfide structure of different thionin types. The cysteine residues are numbered from the N terminus. Cysteine residues that are absent are indicated (\times). The numbers of basic (+) and acidic (-) residues of each thionin type are written under each structure. (b) Comparison of genes *ThV* from wheat (type V) and *HthA* from barley (type I). The numbers express, as percentages, matches/(matches + mismatches) for each of the indicated domains. (c) Average ratios of non-synonymous (kA) to synonymous (kS) nucleotide substitution rates calculated according to the method of Li *et al.* (1984). Shaded bars correspond to binary combinations involving type-V thionin and open bars to the remaining ones. Standard errors of the mean are also represented. Type-I sequences used were those of α and β -thionin from barley (Hernández-Lucas *et al.*, 1986; Ponz *et al.*, 1986), and $\alpha 1$ and $\alpha 2$ -thionin from wheat (A. Castagnaro & C. Maraña, unpublished results). Type-II sequences were G1 and G2 (Gausung, 1987), and DG3, DC4 and DB4 from barley leaves (Bohlman & Apel, 1987).

with the neutral theory of evolution has also been proposed (Graur, 1985; Graur & Li, 1988). According to this alternative, the relative mutability of a given sequence depends on the proportions of certain amino acid residues, rather than on their specific positions, and can be predicted using an empirical index of mutability, Im (Graur, 1985). In the present case, the Im value is higher for type V ($Im = 2.693$) than for type I ($Im = 0.946$) or type II ($Im = 1.505$), which would tend to support a neutralist explanation for the high rate of non-synonymous substitutions. However, it is to be noted that certain features of the type-V thionin, such as six cysteine, two basic and two acidic

residues, have also been acquired by type-IV thionins, at different positions and through a different evolutionary route, but presumably starting from the same ancestral gene. This suggests that positive selection must have played an important role.

Technical assistance from D. Lamóneda and J. García is gratefully acknowledged. The work was supported by grant BIO90-0084 and the Programa PETRI from the Comisión Interministerial de Ciencia y Tecnología.

References

- Bachman, B., Lüke, W. & Hunsmann, G. (1990). Improvement of PCR amplified DNA sequencing with the aid of detergents. *Nucl. Acids Res.* **18**, 1309.
- Balls, A. K., Hale, W. S. & Harris, T. H. (1942a). A crystalline protein obtained from a lipoprotein of wheat flour. *Cereal Chem.* **19**, 279-288.
- Balls, A. K., Hale, W. S. & Harris, T. H. (1942b). Further observations on a crystalline wheat protein. *Cereal Chem.* **19**, 840-844.
- Bohlman, H. & Apel, K. (1987). Isolation and characterization of cDNAs coding for leaf-specific thionins closely related to the endosperm-specific hordothionin of barley (*Hordeum vulgare* L.). *Mol. Gen. Genet.* **207**, 446-454.
- Bohlman, H., Clausen, S., Behnke, S., Giese, H., Hiller, C., Reiman-Philipp, U., Schrader, G., Barkholt, C. & Apel, K. (1988). Leaf-specific thionins of barley: a novel class of cell wall proteins toxic to plant-pathogenic fungi and possibly involved in the defense mechanism of plants. *EMBO J.* **7**, 1559-1565.
- Creighton, T. E. & Charles, I. G. (1987). Sequences of the genes and polypeptide precursors for two bovine protease inhibitors. *J. Mol. Biol.* **194**, 11-12.
- Ebrahim-Nesbat, F., Behnke, S., Kleinhofs, A. & Apel, K. (1989). Cultivar-related differences in the distribution of cell-wall-bound thionins in compatible and incompatible interactions between barley and powdery mildew. *Planta*, **179**, 203-210.
- Feng, D.-F. & Doolittle, R. F. (1987). Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* **25**, 351-360.
- Fernández de Caleyá, R., Hernández-Lucas, C., Carbonero, P. & García-Olmedo, F. (1976). Gene expression in allopolyploids: genetic control of lipopurothionins in wheat. *Genetics*, **83**, 687-699.
- Fioretti, E., Iacopino, G., Angeletti, M., Barra, D., Bossa, F. & Ascoli, F. (1985). Primary structure and anti-proteolytic activity of a Kunitz-type inhibitor from bovine spleen. *J. Biol. Chem.* **260**, 11451-11455.
- Fisher, R., Behnke, S. & Apel, K. (1989). The effect of chemical stress on the polypeptide composition of the intercellular fluid of barley leaves. *Planta*, **178**, 61-68.
- García-Olmedo, F., Rodríguez-Palenzuela, P., Hernández-Lucas, C., Ponz, F., Maraña, C., Carmona, M. J., López-Fando, J., Fernández, J. A. & Carbonero, P. (1989). The thionins: a protein family that includes purothionins, viscotoxins and crambins. In *Oxford Surveys of Plant Molecular and Cell Biology* (Mifflin, B. ed.), vol. 6, pp. 31-60, Oxford University Press.
- García-Olmedo, F., Carmona, M. J., López-Fando, J. J., Fernández, J. A., Castagnaro, A., Molina, A., Hernández-Lucas, C. & Carbonero, P. (1990). Characterizations and analysis of thionin genes. In

- Plant Gene Series: Genes Involved in Plant Defense* (Boller, T. & Meins, F. eds), Springer-Verlag, Heidelberg, in the press.
- García-Olmedo, F., Salcedo, G., Sánchez-Monge, R., Hernández-Lucas, C., Carmona, M. J., López-Fando, J. J., Fernández, J. A., Gómez, L., Royo, J., García-Maroto, F., Castagnaro, A. & Carbonero, P. (1991). Alpha-amylase/trypsin inhibitors and thionins. Possible defense proteins from barley. In *Barley: Genetics, Biochemistry, Molecular Biology and Biotechnology* (Shewry, P. ed.) CAB International, in the press.
- Gausung, K. (1987). Thionin genes specifically expressed in barley leaves. *Planta*, **171**, 241–246.
- Graur, D. (1985). Amino acid composition and the evolutionary rates of protein coding genes. *J. Mol. Evol.* **22**, 53–62.
- Graur, D. & Li, W.-H. (1988). Evolution of protein inhibitors of serine proteinase: positive Darwinian selection or compositional effects? *J. Mol. Evol.* **28**, 131–135.
- Hernández-Lucas, C., Royo, J., Paz-Ares, J., Ponz, F., García-Olmedo, F. & Carbonero, P. (1986). Polyadenylation site heterogeneity in mRNA encoding the precursor of the barley toxin α -hordothionin. *FEBS Letters*, **200**, 103–105.
- Hill, R. E. & Hastie, N. D. (1987). Accelerated evolution in the reactive centre regions of serine protease inhibitors. *Nature (London)*, **326**, 96–99.
- Laskowski, M., Kato, I., Kohr, W. J., Park, S. J., Tashiro, M. & Whatley, H. E. (1987a). Positive Darwinian selection in evolution of protein inhibitors of serine proteinases. *Cold Spring Harbor Symp. Quant. Biol.* **52**, 201–207.
- Laskowski, M., Kato, I., Ardelt, W., Cool, J., Deuton, A., Empie, M. W., Kohz, W. J., Park, S. J., Schateley, B. L., Schoenberger, O. L., Tashiro, M., Vichot, G., Whatley, H. E., Wieczorek, A. & Wieczorek, M. (1987b). Ovomucoid third domains from 100 avian species: isolations, sequences, and hypervariability of enzyme-inhibitor contact residues. *Biochemistry*, **26**, 202–221.
- Li, W.-H., Wu, C. I. & Luo, C. C. (1984). Nonrandomness of point mutation as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. *J. Mol. Evol.* **21**, 58–71.
- Maraña, C., García-Olmedo, F. & Carbonero, P. (1988). Linked sucrose synthase genes in group-7 chromosomes in hexaploid wheat (*Triticum aestivum* L.). *Gene*, **63**, 253–260.
- Murray, M. G. & Thompson, W. F. (1980). Rapid isolations of high molecular weight plant DNA. *Nucl. Acids Res.* **8**, 4321–4325.
- Ponz, F., Paz-Ares, J., Hernández-Lucas, C., Carbonero, P. & García-Olmedo, F. (1983). Synthesis and processing of thionin precursors in developing endosperm from barley (*Hordeum vulgare* L.). *EMBO J.* **2**, 1035–1040.
- Ponz, F., Paz-Ares, J., Hernández-Lucas, C., García-Olmedo, F. & Carbonero, P. (1986). Cloning and nucleotide sequence of a cDNA encoding the precursor of the barley toxin α -hordothionin. *Eur. J. Biochem.* **156**, 131–135.
- Redman, D. G. & Fisher, N. (1969). Purothionin analogues from barley flour. *J. Sci. Food Agric.* **19**, 651–655.
- Reiman-Philipp, U., Behnke, S., Batschauer, A., Shafer, E. & Apel, K. (1989). The effect of light on the biosynthesis of leaf-specific thionins in barley *Hordeum vulgare*. *Eur. J. Biochem.* **182**, 283–289.
- Rodríguez-Palenzuela, P., Pintor-Toro, J. A., Carbonero, P. & García-Olmedo, F. (1988). Nucleotide sequence and endosperm-specific expression of the structural gene for the toxin α -hordothionin in barley (*Hordeum vulgare*). *Gene*, **70**, 271–281.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual* (2nd edit.). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Samuelsson, G. (1974). Mistletoe toxins. *Syst. Zool.* **22**, 566–569.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with the chain-terminating inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.
- Teeter, M. M., Mazer, J. A. & L'Italien, J. J. (1981). Primary structure of the hydrophobic plant protein crambin. *Biochemistry*, **20**, 5437–5443.
- Thomas, P. S. (1983). Hybridization of denatured RNA transferred or dotted to nitrocellulose paper. *Methods Enzymol.* **100**, 18–29.
- Vernon, L. P., Evett, G. E., Zeikus, R. D. & Gray, W. R. (1985). A toxic thionin from *Pyricularia pubera*: purification, properties and amino acid sequence. *Arch. Biochem. Biophys.* **238**, 18–19.

Edited by N.-H. Chua