Available online at www.sciencedirect.com
SCIENCE DIRECT®
Plant Physiology and Biochemistry 42 (2004) 355–365
Plant Physiology and Biochemistry
www.elsevier.com/locate/plaphy

Mini review

The eight-cysteine motif, a versatile structure in plant proteins

Matilde José-Estanyol [a,*], F. Xavier Gomis-Rüth [b], Pere Puigdomènech [a]

[a] Laboratori de Genètica Molecular i Vegetal, Consorci CSIC-IRTA, Jordi Girona 18-26, 08034 Barcelona, Spain
[b] Institut de Biologia Molecular de Barcelona, IBMB-CSIC, Jordi Girona 18-26, 08034 Barcelona, Spain

Received 8 January 2004; accepted 22 March 2004

Available online 20 April 2004

Abstract

A number of protein sequences deduced from the molecular analysis of plant cDNA or genomic libraries can be grouped in relation to a defined number of cysteine residues located in distinct positions of their sequences. This is the case for a group of around 500 polypeptides from different species that contain a small domain (less than 100 amino acids residues) displaying a pattern of eight-cysteines in a specific order. The plant sequences containing this motif belong to proteins having different functions, ranging from storage, protection, enzyme inhibition and lipid transfer, to cell wall structure. The eight-cysteine motif (8CM) appears to be a structural scaffold of conserved helical regions connected by variable loops, as observed by three-dimensional structure analysis. It is proposed that the cysteine residues would form a network of disulfide bridges necessary, for the maintenance of the tertiary structure of the molecule together with the central helical core, while the variable loops would provide the sequences required for the specific functions of the proteins.

Keywords: Eight-cysteine motif; HyPRPs; LTPs; 2S-albumins; Cereal protease inhibitors

1. Introduction

An increasing number of proteins have been characterized during the last years by protein purification, cDNA or genomic cloning and genomic sequencing from different plant species that show similarity in a region containing eight-cysteine residues present in a specific pattern. This protein domain is usually formed by less than 100 amino acids. Sometimes this eight-cysteine motif (8CM) domain is preceded or interrupted by a proline or proline/glutamine repeat domain. The eight-cysteine residues are located following a conserved pattern, where the third and fourth cysteines are consecutive in the polypeptide chain and the fifth and sixth cysteines are separated by only one residue.

The plant sequences containing this motif belong to proteins that share a signal peptide and have different functions. They are largely distributed in the plant kingdom [34,61–62]. A number of these proteins, known as 2S-albumins, have been found in intracellular storage protein bodies of embryo organs. They appear to be very stable glutamine-rich proteins, having a storage function of nitrogen and sulfur, but they are also protease inhibitors and they display antifungal activity [2]. Other proteins are present in the endosperm having specific inhibitory functions against glycolytic or proteolytic enzymes, such as amylases, trypsin and some human plasma serine proteases [14,23,57,59]. Other 8CM sequences belong to a subfamily of cell wall proline-rich structural proteins named hybrid proline-rich proteins (HyPRP) [37], because they mostly contain a proline-rich domain in the N-terminal half after the signal peptide and a hydrophobic 8CM domain in the C-terminal half. These genes are usually expressed in a tissue specific manner or induced by specific stresses or hormones but their function has not been clearly determined [37]. The same type of 8CM domain but with a hydrophilic character can be found in other polypeptides such as in lipid transfer proteins (LTP) [39]. LTPs are tryptophan-depleted proteins that appear to be involved in plant biotic and abiotic stresses [22] and they have been shown to be able to transfer lipids from one membrane to another "in vitro" [39]. They have been located in the cell wall, and it has been proposed that they could be involved in different functions such as: furnishing cutin monomers to the cuticle [66,73], permeabilization of membranes during the antifungal response acting either directly or synergistically with other proteins such as cysteine-rich thionins [7,12,39,71] in a similar manner as 2S-albumins and puroindolines [13,69], by binding to specific membrane lipopro-

* Corresponding author.
E-mail address: mjegmp@ibmb.csic.es (M. José-Estanyol).

doi:10.1016/j.plaphy.2004.03.009

teins or to elicitin membrane receptor sites [5]. Some members of these 8CM protein families are known allergens by food ingestion [49] or involved in asthma processes by inhalation [29,53,60].

2. Alignment of selected protein sequences

A selection of the proteins that belong to the 8CM family aligned to show the pattern of the cysteine residues is shown in Fig. 1. At the present stage of sequences available in the Data Banks, we have found by FASTA, BLAST and INTERPRO (IPR003612); PFAM (PF00234) analysis a minimum of 500 different proteins having this pattern in the databases (Table 1). Sixty-nine sequences (from 29 different species) belong to the HyPRPs [37,46] and around 300 sequences among them (from 71 different species) belong to the LTPs [77]. A collection of 47 proteins (from eight different cereals) belongs to the cereal inhibitor family [67] and at least 81 sequences (from 30 different species) have been identified as part of the 2S-albumin seed storage family of proteins [2,61]. All the 8CM protein families are distributed in both dicot and monocot plant species, although 2S-albumins in monocots have been only reported in rice [64] and maize [79]. One exception is that produced by 8CM protease inhibitors, which is only present in monocotyledonous plants. In dicotyledonous species, these inhibitor functions are associated with other families of proteins having different cysteine patterns, as described below.

The sequencing of the genome of *Arabidopsis thaliana* has allowed us to identify the 8CM containing proteins in a simple plant. Twenty-three different *Arabidopsis* HyPRPs have been identified because they have an 8CM hydrophobic domain placed at the C-terminal end of the protein after a proline-rich domain of variable length preceded by a signal peptide (Table 1). They can be distributed in three groups in relation to the length of their proline-rich domain (long, short, and very short or null). Fourteen of these sequences are present in chromosome IV, and four, two, one and two, in chromosomes I–III and V, respectively. One of these proteins has previously been described as pEARLI1 an aluminum responsive gene [56]. Seven consecutive copies of pEARLI1 can be found in chromosome IV and two in chromosome I.

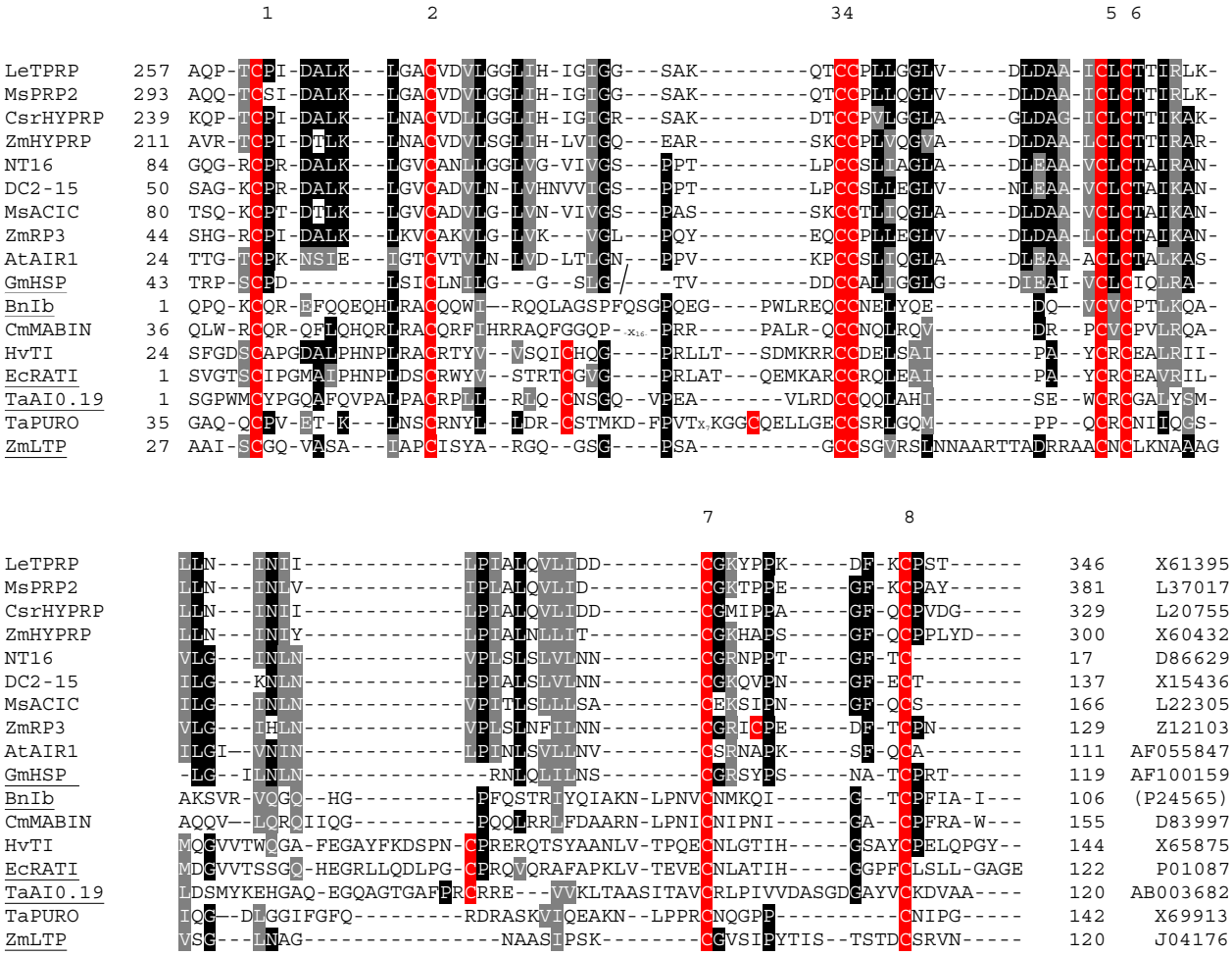

Fig. 1. Clustal alignments of 8CM proteins. Clustal analysis was obtained by alignment of 8CM proteins from http://searchlauncher.bcm.tmc.edu and http://www.ch.embnet.org. Protein names and accession numbers are indicated. Numbers at the top of alignments indicate the position of each cysteine in the motif. Residues are colored according to similarity group. Cysteines are red colored. Proteins with know tertiary structure are underlined. (X_{16}) Gap of 16 non-related aminoacids. (/) End and beginning of light and heavy chain, respectively.

Table 1
CM8 containing proteins obtained by BLAST queries in the EMBL Genebank http://www2.ebi.ac.uk and *Arabidopsis* genome MIPS http://mips.gsf.de from zmHyPRP for HyPRPs [38], AtLTP1 for LTPs [72], AtALB1 for 2S-albumins [32] and RATI for cereal inhibitors [67]

Group	Organism	Number
Proteins with an hydrophobic 8CM domain (HyPRP,Hy GRP, others)	*Arabidopsis*	
	At1g12090; At1g12100; At1g62500; At1g62510; At2g10940; At2g45180; At3g22120; At4g12470; At4g12480; At4g12490; At4g12500; At4g12510; At4g12520; At4g12530; At4g12550; At4g15160; At4g22460; At4g22470; At4g22490; At4g22520; At4g22610; At5g46890; At5g46900	23
	Other species	
	Asparagus officinalis X82413; *Brassica napus* X94976; X71618; *Catharanthus roseus* Z26880; X85206; *Cicer arietinum* AJ278505; *Citrus junos* AY100448; *Cucumis sativus* AF104392; *Cuscuta reflexa* L20755; *Daucus carota* X15436; AB000505; *Fragaria x ananassa* AF026382; *Glycine max* AF100159; X69640; *Lithospermum erythrorhizon* D45901; *Lycopersicon esculentum* X61395; X95262; *Malus domestica* U80271; *Medicago sativa* L22305; L37017; *Medicago truncatula* Y15372; *Nicotiana glauca x Nicotiana langsdorffii* D26454; *Nicotiana tabacum* D86629; AF043554; L13439; L13440; L13441; L13442; L13443; *Oryza sativa* L27208; L27209; L27210; *Phaseolus vulgaris* U34333; *Pinus taeda* AF101789; *Pisum sativum* X67427; *Populus nigra* D83227; *Solanum brevidens* U30304; *Solanum melongena* AB032755; *Triticum aestivum* U73214; *Vitis riparia* AF220197; *Vitis vinifera* AY046416; *Zea mays* X60432; Z12103; AF001634; AB018587; AB018588	46
LTPs	*Arabidopsis*	
	At2g15050; At2g18370; At2g38530; At2g38540; At3g08770; At3g51590; At3g51600; At4g33355; At5g01870; At5g59310; At5g59320	11
	Other species	
	Aerides japonica AF198168; *Amaranthus caudatus* P80450; *Avicennia marina* AF331710; *Atriplex nummularia* AB112477; *Beta vulgaris* X92748; *Brassica napus* U22105; U22174; U22175; X60318; AJ245873; *Brassica oleracea* L33904; L33905; L33906; L33907; L29767; AF093751; *Brassica rapa* L31938; *Bromus inermis* AY057932; *Capsicum annuum* AF118131; AF208833; AF208834; *Cicer arietinum* AJ002958; *Citrus sinensis* AF369931; *Corylus avellana* AF329829; *Daucus carota* M64746; *Davidia involucrata* AY059472; *Eleusine coracana* P23802; *Euphorbia lagascae* AF363505; *Fragaria x ananassa* AJ315844; *Gerbera hybrida* Z31588; *Glycine max* AI748411; AW152885; *Gossypium barbadense* AF531366; *Gossypium hirsutum* S78173; U64874; U15153; AF044204; AF195863; AF195864; AF195865; AF228333; *Helianthus annuus* X92648; *Hevea brasiliensis* AY057860; *Hordeum vulgare* X59253; X60292; X68654; X68655; X68656; Z66528; Z66529; Z37114; Z37115; U18127; U63993; X96979; *Lilium longiflorum* AF171094; *Lycopersicon esculentum* X56040; U81996; *Lycopersicon pennellii* U66466; U66465; *Malus x domestica* AF221502; *Nicotiana glauca* AF151214; *Nicotiana tabacum* AF519812; D13952; X62395; *Oryza sativa* D15364; D15678; D22795; D16036; U77295; Z23271; U29176; X83433; X83434; X83435; U31766; AF017358; AF017359; AF017360; AF017361; AF114829; Y08691; AY335485; AF051369; *Pachyphytum* sp L14770; *Phaseolus aureus* P83434; *Phaseolus vulgaris* U72765; *Prunus armeniaca* P81651; *Prunus avium* AF221501; *Prunus domestica* P82534; *Prunus dulcis* (*amygdalus*) X96714; X96716; X96715; *Prunus persica* AJ277163; P81402; AY093700; *Pyrus communis* AF221503; *Retama raetam* AF439280; *Ricinus communis* P10973; P10974; *Setaria italica* AF439446; *Solanum tuberosum* AF525362; AF525363; *Sorghum bicolor* X71667; X71668; X71669; *Spinacia oleracea* M58635; *Triticum aestivum* X05168; AF302788; AF334185; Q9S876; Q9S877; AF551849; AY226580; AF302788; *Triticum turgidum* (*durum*) X63669; *Vitis vinifera* AF467945; AF467946; *Vitis berlandieri x Vitis vinifera* AF465408; *Zea mays* M57249; S45635; U66105; J04176	123
LTPs-like GPI-anchored	*Arabidopsis*	
	At1g18280; At1g73890; At2g13830; At2g27130; At2g44290; At2g44300; At2g48130; At3g22600; At3g22611; At3g43720; At3g58550; At4g08670; At4g12360; At4g22630; At4g22640; At5g09370; At5g13900; At5g64080	18
LTPs-like	*Arabidopsis*	
	At1g27950; At1g32280; At1g36150; At1g43665; At1g43668; At1g48750; At1g52415; At1g55260; At1g62790; At1g64240; At1g66850; At1g70240; At1g73550; At1g73560; At1g73780; At2g13820; At2g15325; At2g37870; At2g48140; At3g07450; At3g12545; At3g18280; At3g22570; At3g22580; At3g53980; At3g57310; At4g08530; At4g14815; At4g22610; At4g22650; At4g28395; At4g30880; At4g33550; At5g05960; At5g38160; At5g38170; At5g38180; At5g38195; At5g48485; At5g48490; At5g52160; At5g55410; At5g55450; At5g55460; At5g56480; At5g62080. Anther specific: At5g07230	47
	Other species	
	Allium cepa S79815; *Ambrosia artemisiifolia* U89793; *Ananas comosus* AY098530; *Brassica rapa* AB010433; *Hordeum vulgare* AF039024; X15257; X69793; U88090; X56547; *Medicago truncatula* Y15371; *Nicotiana tabacum* AF233297; U14167; U14168; *Parietaria judaica* X85012; X95865; X95866; X95867; *Pinus radiata* AF110332; U90342; *Pinus taeda* U10432; *Phaseolus vulgaris* U34334; *Oryza sativa* A23332; P83210; U16721; *Ricinus communis* M86353; M86354; D11077; *Senecio odorus* L33791; *Silene latifolia* Y08779; *Striga hermonthica* Y16247; *Triticum aestivum* P82900; P82901; P39085; *Triticum turgidum* AJ297768; *Vigna unguiculata* X79604; *Zea mays* P83506; *Zinnia elegans* U19266; Anther specific: *Brassica campestris* AY237725; *Lycopersicon esculentum* Z14088; *Lilium henryii* X80718; X80719; *Lilium longiflorum* D21807; D21808; D21809; *Oryza sativa* D50575; *Pinus radiata* U90343; U90350; *Silene latifolia* Y08780; *Zea mays* AJ224355; AJ006702; Puroindolines a: *Avena sativa* AJ249930; *Hordeum vulgare* AJ249929; *Secale cereale* AJ249932; *Triticum aestivum* X69913; X69914; *Triticum monococcum* AJ249933; *Triticum tauschii* AJ249935; Puroindolines b: *Avena sativa* AJ249931; *Hordeum vulgare* AJ249928; *Triticum aestivum* X69912; *Triticum tauschii* AJ249936; *Triticum monococcum* AJ249934; Grain softness: *Aegilops tauschii* AF177219; AY252046; AY252062; AY252063; AY252046; *Triticum aestivum* AF177218; AF177219; AY255771; S72696; X80379; X80381; *Triticum monococcum* AJ242717	74

(continued on next page)

Table 1
(continued)

Group	Organism	Number
2S-albumins	*Arabidopsis*	
	At4g27140; At4g27150; At4g27160; At4g27170; At5g54740	5
	Other species	
	Anacardium occidentale AY081853; *Arachis hypogaea* AF091737; AF092846; AF366561; AY007229; AY117434; AY158467; *Bertholletia excelsa* AB044391; X54490; X54491; *Brassica carinata* X74813; *Brassica juncea* P80207; X65972; *Brassica napus* AF448054; P17333; P09893; K01544; J02782; P24565; X14492; X17542; X58142; K01545; J02586; U04944; *Brassica nigra* X65971; *Brassica oleracea* X65038; X65970; *Brassica rapa* (*campestris*) M64631; M64632; X65969; *Capparis masaikai* P80351; P30233; P80352; P80353; *Carya Illinoinensis* AY192569; *Cucurbita* cv. D16560; *Glycine max* AF005030; U71194; *Gossypium hirsutum* M86213; M83301; *Helianthus annuus* AJ275962; X06410; X76101; *Juglans nigra* AY102930; *Juglans regia* U66866; *Linum usitatissimum* AJ414732; AJ414733; *Lupinus angustifolius* X53523; *Momordica charantia* AJ488931; *Oryza sativa* X63990; D50643; L12252; *Picea glauca* X63193; AF074937; AF074938; AF074939; L47745; U92077; *Pinus strobus* X62433; X62434; X62435; X62436; *Pseudotsuga menziesii* AF029970; AF029972; AF029973; *Raphanus sativus* M63841; M63842; M63843; *Ricinus communis* X54158; *Sesamum indicum* AF091841; AF240005; *Sinapis alba* S54101; X91798; X91799; X91800; X91801; X91802; *Vitis vinifera* AY267254; AY267255; *Zea mays* AF371278	81
Amylase inhibitors	*Arabidopsis*	
	None	
	Other species	
	Monomeric: *Hordeum vulgare* X63517; *Sorghum bicolor* (*vulgare*) P81367; P81368; *Triticum aestivum* P01083	35
	Dimeric: *Hordeum vulgare* AJ009801; *Triticum aestivum* P01084; P01085; *Triticum turgidum* (*durum*) X55454; X16733	
	Tetrameric: *Hordeum vulgare* X69937; X69938; X69939; X69911; *Triticum aestivum* X59791; X17573; X17574; X17575	
	Non-determined: *Hordeum vulgare* X13433; *Triticum turgidum* (*durum*) X16733; X55454; *Oryza sativa* X66257; D11433; D11432; D11430; D11434; D42139; D42140; D42141; D43657; D43658; D43659; X62091; O49178; AP005197; Q8H4MS; Q8H4M4	
Trypsin inhibitors	*Arabidopsis*	
	None	
	Other species	
	Hordeum spontaneum AJ222974; AJ222975; *Hordeum vulgare* AJ222977; AJ223458; AJ222978; X98593; X98594; Y12069	8
Trypsin inhibitors with different inhibitory functions	*Arabidopsis*	
	None	
	Other species	
	Zea mays X54064 (trypsin, insect α-amylase, Hageman factor inhibitor); *Eleusine coracana* P01087 (mammalian, insect α-amylases and trypsin inhibitor); *Hordeum vulgare* X65875 (trypsin, Hageman Factor, kallikrein); AJ251931 (trypsin, insect α-amylase)	4

Identities range from 90% for three of the seven copies in chromosome IV, to 75% for the other three copies in chromosome IV and the two in chromosome I and 60–70% for their two relatives in chromosome V, for the one in chromosome II and for AIR1 one of the *Arabidopsis* HyPRPs with null proline-rich domain whose expression is related to lateral root formation induction by auxin [46]. Between the sequences sharing a long proline-rich repetitive domain are the genes atcwlp (chr. III) and atcwlp-h (chr. IV) [30] cloned by a yeast based signal sequence trap selection method, from an *Arabidopsis* cDNA library, to find secreted and plasma membrane proteins.

Arabidopsis LTP1 [72] sequence homology studies indicate that the first 11 sequences to come out from the search correspond to the classical LTPs [1,15]. These sequences present the consensus sequence described for maize LTP [63] necessary to establish an internal hydrophobic cavity to allow lipid transfer between membranes. This consensus sequence is progressively lost in the aligment as sequences are more distant from LTP1. In addition to the classical LTPs we have identified 65 "LTP-like" proteins, 29 of them sharing C-terminal ends of varying length, and 18 of them sharing putative GPI-anchors [8,74] (Table 1).

A Fasta and Blast analysis has also been done with AtALB1 protein [32] to search for *Arabidopsis* 2S-albumin seed storage proteins and it has only allowed the identification of five sequences. Four of them (ALB1-ALB4) are present in chromosome IV [32] and the fifth isoform is localized in chromosome V [76] (Table 1).

The phylogenetic analysis of all these proteins shows particularly interesting details. When all HyPRP sequences from either *Arabidopsis* or other plants are plotted in an unrooted tree it is possible to observe different groups that are related either to the length of the proline-rich domain attached to the 8CM, to its specific expression pattern in the root (ZmAF001634, ZmZ12103, OsL27210, OsL27209) or to the presence of glycine instead of proline in the repetitive protein domain (LeX95262, NtD86629, NtAF043354, SmAB032755) (Fig. 2). The opposite situation in the tree, of proteins containing long and short proline-rich repeats, suggests that they are probably the result of repeat amplification or insertion into hydrophobic 8CM proteins present in the central region of the tree, followed by divergent evolution.

The phylogenetic tree analysis of LTPs (not shown) shows three main groups of proteins. Classical LTPs spread in two branches, one for dicotyledonous and one for monocotyledonous plants and LTPs-like proteins spread as a third branch where LTPs sharing a putative GPI anchor at their C-terminal end are localized. LTP-like proteins have a low level of similarity in the amino acids consensus for the hydrophobic

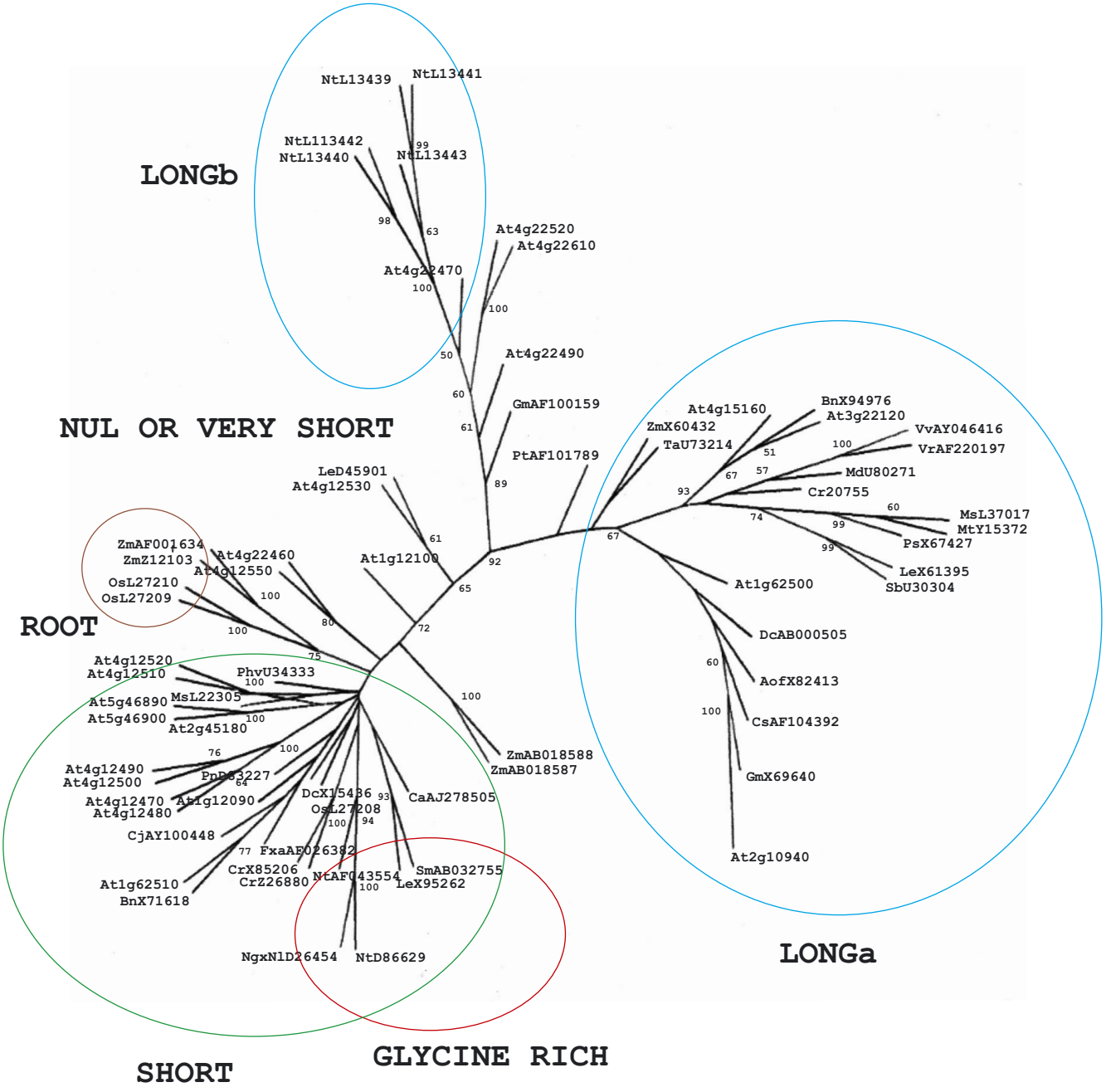

Fig. 2. Phylogenetic tree of plant HyPRPs. Proteins were identified by BLAST queries (EMBL Genebank http://www2.ebi.ac.uk and *Arabidopsis* genome MIPS http://mips.gsf.de) from zmHyPRP 8CM domain [38]. Protein sequences were aligned with the genedoc program (http://www.psc.edu/biomed/genedoc/). The construction of the unrooted phylogenetic tree was carried out using the Phylip's package program [19]. Bootstrap values above 50% are shown.

internal channel, and may have evolved from a different gene that result from a very early division of the ancestral LTP gene in plants [77]. Representative of this group are a potent antimicrobial protein from onion seeds (Ace-AMP1) [12] that has been shown to be unable to transfer lipids and that will be described below and anther specific proteins [41,54]. They also include puroindolines, proteins first described in wheat as grain softness proteins [9,24,36] and that are LTP-like proteins with 10 cysteines and a tryptophan-rich domain. A number of proteins similar to puroindolines have been identified in other plant species [25].

3. Three-dimensional structure of 8CM proteins

Four members representative of the 8CM family of proteins, soybean hydrophobic seed protein (GmHSP) [3], wheat α-amylase dimeric inhibitor (TaAI0.19) [47], bifunctional corn Hageman factor inhibitor (CHFI) [4] and bifunctional α-amylase/trypsin inhibitor from ragi seeds (EcRATI) [31,67] have been crystallized and their tertiary structure is solved. In the four cases the cysteine bridges are those formed between residues 1–5, 2–3, 4–7 and 6–8 (Fig. 3). The same pattern of disulfide bridges has been found by nuclear

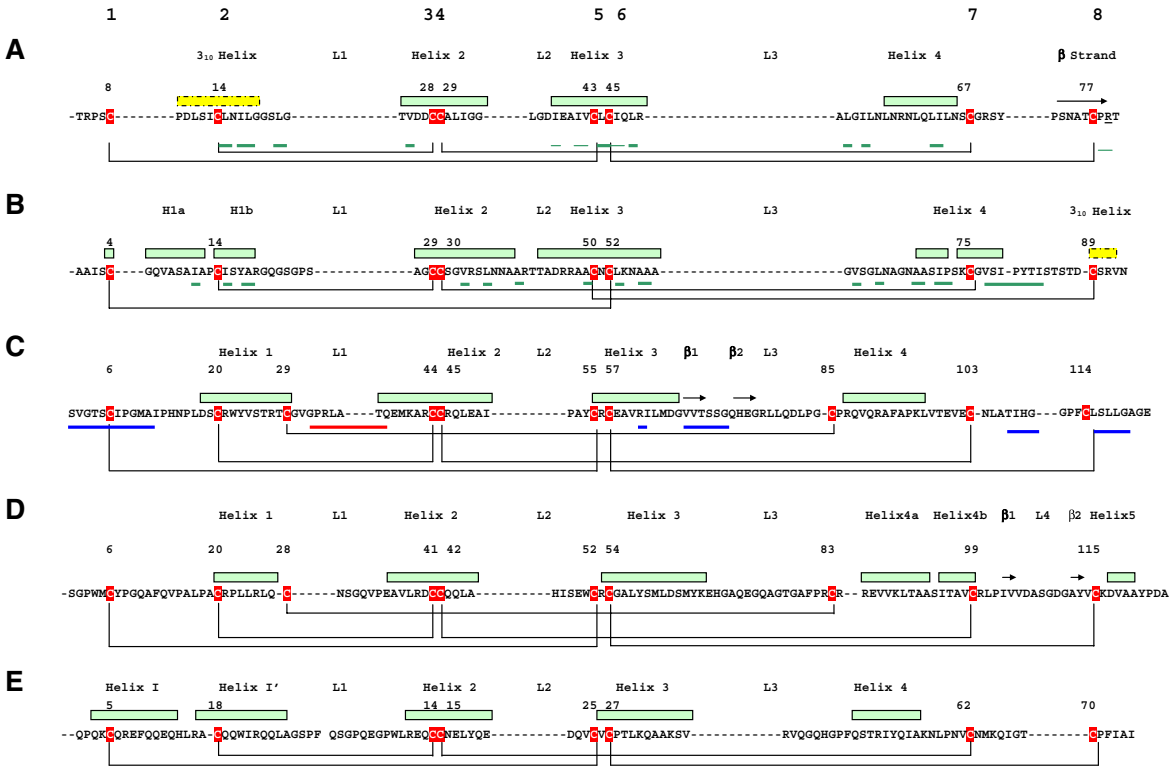

Fig. 3. Schematic representation of the disulfide bond patterns in GmHSP (**A**), ZmLTP (**B**); EcRATI (**C**); HvAI0.19 (**D**) and BnIb (**E**). Above the sequences are the regions of the α-helices (green boxes), β-strands (horizontal arrows), and 3_{10} helix (yellow boxes). L, loop. Binding sites for trypsin and α-amylase are underlined in red and blue, respectively. Aminoacids involved in ZmLTP and ZmHSP hydrophobic cavity are underlined in green (thick and thin lines for ZmHSP correspond to the larger and smaller cleft).

magnetic resonance (NMR) for three albumin genes namely: napin (BnIB) a 2S-albumin from *Brassica napus* purified from *Brassica* seeds [58] and from its recombinant protein [52] expressed in *Pichia pastoris* [50]; an allergen 2S-albumin recombinant protein from *Ricinus communis* (RicC3) [51] expressed using minimal medium cultures of *E. coli* [20] and for SFA-8 albumin (HaSF8) purified from sunflower seeds [52] and from its cyanogen bromide cleavage [17].

For LTPs the tertiary structure was solved in solution by NMR in different species: maize [28], wheat [27], barley [33] and rice [55], and also by X-ray crystallography in maize [63] and rice [43]. Results obtained by NMR or X-ray crystallographic structure studies of LTPs are in agreement. LTP proteins show a difference in the disulfide bridge pattern with respect to the other 8CM proteins involving cysteine residues 5 and 6 that are separated by a single residue (Fig. 3). Nevertheless, all 8CM proteins share a similar tertiary structure and define a new fold class formed by four helices connected by loops of different length (Fig. 3).

The three-dimensional structures of the five most representatives 8CM proteins have been superimposed in Fig. 4A. From these comparisons it is possible to conclude that the position of the C2–C3 and C4–C7 disulfide bonds in the scaffold is conserved for the five proteins. C1–C5 and C6–C8 are similarly positioned in EcRati, TaAI0.19 and BnIb, as a result of the presence of C5 and C6 at the beginning of helix 3. On the other hand, GmHSP C1–C5 and ZmLTP C1–C6

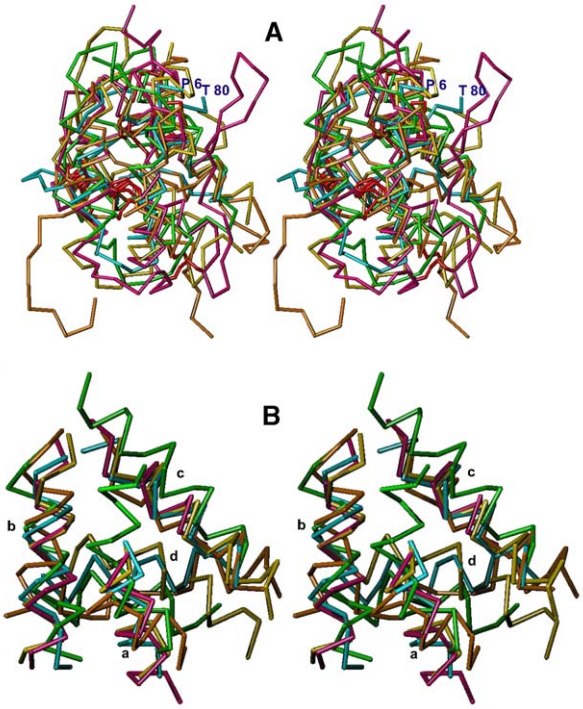

Fig. 4. Superimposition of the tertiary structure of different 8CM proteins. **A**, Stereoview of the superimposition of the average NMR structure of BnIb (brown) and the C^{α} traces of the X-ray structures of Hv AI0.19 (yellow), EcRATI (red), ZmLTP (green) and GmHSP (blue). Red sticks symbolize disulfide bonds. **B**, Same stereoview as in (A) but showing only the α-helices helix 1 (a), helix 2 (b), helix 3 (c), helix 4 (d).

shows topological similarity, distinct from the previous three cases mentioned, but shows different location for their other dicysteine bridge (GmHSP C6–C8 and ZmLTP C5–C8).

The common topology of these structures is inferred from the comparison of their four α-helical structures segments (Fig. 4). Structural similarity between the four superimposed helices of the five studied proteins is shown in Fig. 4A, B to highlight their conservation in the relative distribution (Fig. 4B). Maximum divergence can be observed in helix 4 and in the loops between the helices of the different proteins analyzed. Particularly, interesting is loop 1, connecting helix 1 and helix 2, where the trypsin inhibitor site with its consensus sequence GPRL (one letter amino acid code) is observed in RAGI (Figs. 1 and 4A), as well as in CHFI [4] (Fig. 1). Divergences can also be observed for maize LTP internal hydrophobic cavity and in its C-terminal end that appears to be adapted to locate lipids. Other differences are disruption of helices 1 and 4 in zmLTP and of helix 1 in BnIb and the presence of β-sheets structures after helix 4 in GmHSP and TaAI0.19.

Superimposed stereoviews of 8CM proteins have shown conserved distribution of helices, where cysteine residues form a network of disulfide bridges necessary for the maintenance of the tertiary structure of the molecule. Major differences are located in the loops and N- and C-terminal tails, where the main functional divergences of 8CM structural scaffold are found, as will be described below.

4. Comparison of the three-dimensional structure differences of 8CM proteins with HSP

To better visualize differences between tertiary structures of 8CM proteins, the structures of BnIb [58], TaAI0.19 [47], EcRATI [31] and ZmLTP [63] have been individually superposed to GmHSP [3] (Fig. 5). GmHSP has been chosen as a reference for the comparison for its consensus structure of four helices connected by short loops. GmHSP is a highly hydrophobic HyPRP [48] allergenic protein [29], mainly present in the seed pericarp, where it is supposed to have a structural or defense function [26]. Fig. 5A shows GmHSP [3] superposition with BnIb [58], a 2S-albumin. BnIb tertiary structure presents two polypeptide chains linked by two disulfide bridges as the result of the classical cleavage of 2S-albumins loop 1 by maturation proteolytic enzymes during translation [58], with one exception, the above cited sun8 [17]. Differences with GmHSP are observed in the length of loop 3, in the C-terminal tail orientation and distribution of charged amino acids. Conversely to HSP, BnIb presents 23 mainly positively, surface located, charged amino acids. These charges might be responsible for electrostatic interactions with the membrane during defense processes.

Fig. 5B compares GmHSP [3] with dimeric TaAI0.19, an α-amylase inhibitor from wheat kernel [47], that inhibits α-amylases from various sources, such as human saliva, chick pancreas, yellow mealworm and *Bacillus subtilis* by a not well understood mechanism. TaAI0.19 structure shows a

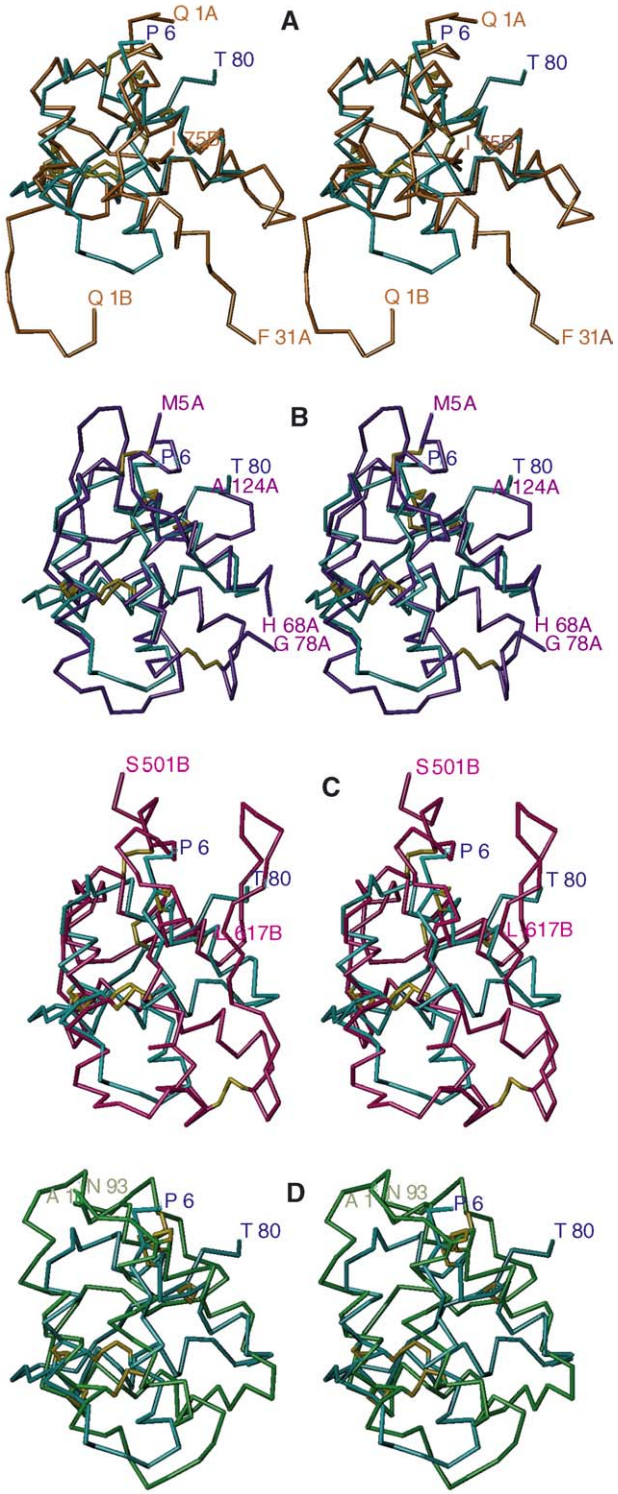

Fig. 5. Superpositions of GmHSP (blue) with (**A**) BnIb (brown), (**B**) HvAI0.19 (purple), (**C**) EcRATI (red), (**D**) ZmLTP (green). Disulfide bonds as yellow stick.

breakage in loop 3 just before the additional disulfide bond placed before helix 4. Although the four α-helices of GmHSP and TaAI 0.19 can be superimposed, the conformation of the segments connecting the α-helices (specially loop 3) differ considerably. In addition, the C-terminal segment differs by two antiparallel β-strands and an additional short helix 5.

GmHSP [3] and EcRati [31], a bifunctional inhibitor that is the prototype of the cereal inhibitor superfamily that further includes trypsin and α-amylases inhibitors from different sources, are compared in Fig. 5C. Differences are observed in the N-terminal tail and loop 1 of EcRati, involved in amylase and trypsin inhibitory functions, respectively, and in loop 3 with an antiparallel β-sheet.

GmHSP [3] and ZmLTP [63] are compared in Fig. 5D. Both proteins show a similar scaffold structure, although the C5 and C6 bridges are inverted in ZmLTP with respect to GmHSP. Differences in the polar cluster of the entrance of the internal hydrophobic cavity and in its size as well as in the C-terminal tail orientation are also observed. These differences may explain the inability of HSP to transfer lipids, although a hydrophobic cleft is observed inside the folded protein, while differences in the surface charge of both proteins explain the insolubility of HSP that has a hydrophobic surface charge.

Functionality of the internal hydrophobic LTP cavity has been probed by crystallographic structure studies of maize LTP complexed with lipids such as palmitate, as well as solution structure studies. Results show that only one acyl chain of the hydrophobic tails of fatty acids is inserted in the protein internal cavity. Different conformational changes are observed after lipid binding to the different LTPs as a consequence of their different cavity size and accessibility [40,44,63,65]. A limit example of variability in the internal cavity is given by Ace-AMP1, a potent antimicrobial LTP-like protein described in onion seeds [69] and that has two tryptophan residues in its primary sequence. Solution studies of Ace-AMP1 show that the presence of aromatic amino acids in the cavity of the folded protein may avoid lipid binding to the protein. However, its interaction with phospholipid membranes appears to be more efficient than the one described for wheat LTP, due to the presence of 19 arginine residues spread over all the sequence, 12 of them being oriented toward the C-terminal region. This structure may be important for the interaction with acidic lipid layers. Nevertheless, this interaction appears not to be so efficient as the one described for puroindolines [13], involved in membrane interactions as albumins and thionins [69].

5. Relation of 8CM proteins with other families of plant proteins

The 8CM domain has shown to be able to detect in databases different groups of proteins with different functions. These functions are also shared by other proteins with different tertiary structures. A good example is provided by the proteins having a trypsin inhibitor function that is done by plant proteins that have very different conformations but have in common the way they interact with proteinases through an exposed binding loop [6]. These trypsin inhibitor families are usually small proteins cross-connected by disulfide bridges associated to β-strand conformations instead of the four α-helices described for the 8CM family of proteins.

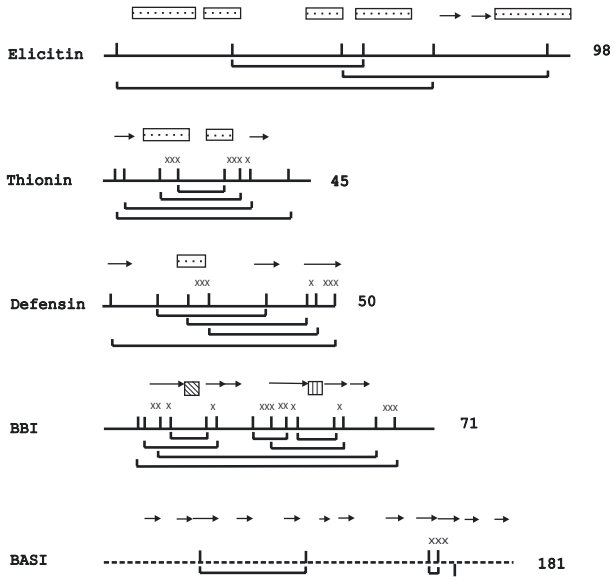

Fig. 6. Schematic representation of the disulfide bond patterns and tertiary structure in *Phytophthora cryptogea* β-cryptogein β-elicitin [18], wheat α1-purothionin [16], *Raphanus sativus* AFP1 defensin [11], BBI (bifunctional soybean Bowman–Birk) [78] and BASI (barley α-amylase-subtilisin inhibitor) [75]. Symbols placed above the protein schemes indicate the protein regions corresponding to α-helix (boxes with points) and β-strand (horizontal arrows) structures. Recognition loop sites for trypsin and chymotrypsin enzymes are indicated by boxes with diagonal and vertical lines, respectively. Cysteine positions are indicated by vertical lines on the protein schemes. BASI Arg 155 involved in inhibition of barley α-amylase isozyme (AMY2) is indicated by a vertical line under BASI schema. Protein amino acid number is indicated at the right of the protein schemes. BASI sequence is represented with a dotted line instead of a continuous one, to indicate that it is represented in a different scale with respect to the other proteins described in the Figure. X makes reference to the number of aminoacids residues present between cysteines in the different cysteine patterns described for each group of proteins.

One example is the Bowman–Birk family of double-headed proteinase inhibitors with 14 cysteine residues embedded in a consensus sequence and associated to a six β-strand protein conformation, which is stabilized by seven disulfide bridges [78] (Fig. 6). Another example is provided by the soybean Kunitz trypsin inhibitor (GmKTI) with two disulfide cysteines bridges (1–2, 3–4) and with loops wrapping around the hydrophobic side chains of β-pleated sheet structure [68].

Sequence homology of GmKTI is observed with other inhibitors, such as BASI (barley bifunctional α-amylase-subtilisin inhibitor) [75] with 12 antiparallel β-strands joined by loops and the same cysteine bridge pattern described for GmKTI (Fig. 6). In this way, the amylase inhibitory function, as it was above seen for trypsin inhibitors, is associated with different structural conformations, but in this case similar regions of the structurally related α-amylases interact with the substrate in a way that appears to be different for the different amylase inhibitors [67].

Another example of different structure is elicitins produced by the pathogenic fungal *Phytophthora* genus. Elicitins are small fungal elicitor proteins able to bind and transfer sterols between artificial membranes, as LTPs do with lipids [45]. Tertiary structure of elicitins is different from the

one described for LTPs. For the β-cryptogein β-elicitin protein, five α-helices and two β-sheets and a Ω loop have been defined forming a hydrophobic cavity of highly conserved residues that could be the plausible binding functional site [18]. In this case, the structure is stabilized by three disulfide bridges (Fig. 6).

Other small cysteine containing proteins involved in plant defense response, as some described 8CM domain proteins are thionins and defensins [10]. Thionins are proteins with six or eight-cysteines, localized in protein bodies and/or cell walls of different species [21], while defensins have eight- or 10-cysteines and are localized in vacuoles or cell walls [11]. In both cases the mature proteins have a half of the amino acids of the 8CM protein domain. Thionin tertiary structure as described for α1-purothionin [16] results in a hydrophobic underside with a structure of a Greek gamma letter, where the vertical stem is a pair of α-helices and the horizontal arm are a strand and short antiparallel β-sheet stabilized by four disulfide bridges (Fig. 6) in the core of the protein [16]. Defensins have adopted a different conformation [11], with a triple stranded antiparallel β-sheet, and a single α-helix lying in parallel with the β-sheet stabilized by four [42] or by five bridges, as run in PhD1-2 defensin where the two additional cysteines form a new cysteine bond [35,42](Fig. 6).

A different case is that of some cereal storage prolamins that have a modified 8CM pattern of cysteines either by the addition or deletion of these residues. The structure of a prolamin, a wheat γ-gliadin, has been studied by circular dichroism of the full protein and of two peptides corresponding to the repetitive and non-repetitive domain indicating a structure of α-helices stabilized by disulfide bridges in a pattern different from the one described for 8CM proteins [70]. It has been proposed that prolamins may have the same evolutionary origin that 8CM proteins do [34,61–62]. For all these proteins, three common regions A, B and C have been described in the non-repetitive protein domain. The same regions have been shown to be present in 2S-albumins and 8CM protease inhibitors, also described herein [34,61–62]. As these regions show sequence similarity with each other, it has been proposed that they may have evolved from a single short ancestral sequence [62]. Correspondence between A–C prolamin regions and the 8CM family described here is deduced from their comparison with EcRATI and HvTI (Fig. 3) [62].

6. Conclusion

The presence of the 8CM in a large family of proteins (around 500 sequences currently present in databases) having very different functions allows to propose that this is a motif that has essentially a structural and evolutionary meaning, but that does not bear in itself a functional role. The conserved structure includes four disulfide bridges and four α-helices as observed in the proteins whose 3D-structure has been solved so far. It is also possible to conclude that 8CM proteins have a common ancestor with other proteins, such as cereal prolamins, whose structure has significantly diverged during the time course of evolution. Based on this 8CM protein scaffold, the regions between the cysteine residues would have evolved independently providing the necessary flexibility to accommodate different functionalities. We propose that the 8CM proteins have evolved in different directions as a consequence of a set of different variations; for instance, by the appearance of an internal hydrophobic cavity adapted to lipid membrane transfer that results in LTPs, by the addition of a proline-rich domain to a hydrophobic 8CM in the case of HyPRPs or of a proline/glutamine-rich domain in prolamin proteins, by the introduction of inhibitory elements in specific loops between helices and protein terminal ends in cereal inhibitors, or by the introduction of sequences that promote loop proteolytic cleavage in 2S-albumin families. These modifications have resulted in contributions to the defense program of the seed, in providing surface cell markers or in developing important allergens by not a well understood scaffold modification, surface positive charge or local glycosilation. Therefore, the 8CM appears as a common and frequently used structure in plant proteins, a framework that may serve to evolve different functions from a single sequence scaffold by adding, in a combinatorial way, specific functional elements.

Acknowledgements

The work was supported by grant Bio1997-0729 and Bio2003-00132 from Plan Nacional de Investigación Científica y Técnica and within the framework of Center de Referència de Biotecnologia de la Generalitat de Catalunya.

References

[1] V. Arondel, Ch. Vergnolle, C. Cantrel, J.-C. Kader, Lipid transfer proteins are encoded by a small multigene family in *Arabidopsis thaliana*, Plant Sci. 157 (2000) 1–12.

[2] J. Barciszewski, M. Szymanski, T. Haertlé, Minireview: analysis of rape seed napin structure and potential roles of the storage protein, J. Protein Chem. 19 (2000) 249–254.

[3] F. Baud, E. Pebay-Peyroula, C. Cohen-Addad, S. Odani, M.S. Lehmann, Crystal structure of hydrophobic protein from soybean; a member of a new cysteine-rich family, J. Mol. Biol. 231 (1993) 877–887.

[4] C.A. Behnke, V.C. Yee, I. Le Trong, L.C. Pedersen, R.E. Stenkamp, S.-S. Kim, G.R. Reeck, D.C. Teller, Structural determinants of the bifunctional corn Hageman Factor inhibitor: X-ray crystal structure at 1.95 Å resolution, Biochemistry 37 (1998) 15277–15288.

[5] J.P. Blein, P. Coutos-Thévenot, D. Marion, M. Ponchet, From elicitins to lipid-transfer proteins: a new insight in cell signalling involved in plant defence mechanisms, Trends Plant Sci. 7 (2002) 293–296.

[6] W. Bode, R. Huber, Structural basis of the endoproteinase–protein inhibitor interaction, Biochim. Biophys. Acta 1447 (2000) 241–252.

[7] H. Bohlmann, S. Clausen, S. Behnke, H. Giese, C. Hiller, U. Reimann-Philipp, G. Schrader, V. Barkholt, K. Apel, Leaf-specific thionins of barley a novel class of cell wall proteins toxic to plant-pathogenic fungi and possibly involved in the defence mechanism of plants, EMBO J. 7 (1988) 1559–1565.

[8] G.H.H. Borner, J. Sherrier, T.J. Stevens, I.T. Arkin, P. Dupree, Prediction of glycosylphoshatidylinositol-anchored proteins in *Arabidopsis*. A genomic analysis, Plant Physiol. 129 (2002) 486–499.

[9] G. Branlard, N. Amiour, G. Igrejas, T. Gaborit, S. Herbette, M. Dardevet, D. Marion, Diversity of puroindolines as revealed by two-dimensional electrophoresis, Proteomics 3 (2003) 168–174.
[10] W.F. Broekaert, B.P.A. Cammue, M.F.C. De Bolle, K. Thevissen, G.W. De Samblanx, R.W. Osborn, Antimicrobial peptides from plants, Crit. Rev. Plant Sci. 16 (1997) 297–323.
[11] W.F. Broekaert, F.R.G. Terras, B.P.A. Cammue, R.W. Osborn, Plant defensins: novel antimicrobial peptides as components of the host defense system, Plant Physiol. 108 (1995) 1353–1358.
[12] B.P.A. Cammue, K. Thevissen, M. Hendriks, K. Eggermont, I.J. Goderis, P. Proost, J. Van Damme, R.W. Osborn, F. Guerbette, J.-C. Kader, W.F. Broekart, A potent antimicrobial protein from onion seeds showing sequence homology to plant lipid transfer proteins, Plant Physiol. 109 (1995) 445–455.
[13] P. Charnet, G. Molle, D. Marion, M. Rousset, V. Lullien-Pellerin, Puroindolines form ion channels in biological membranes, Biophys. J. 84 (2003) 2416–2426.
[14] G.L. Chong, G.R. Reeck, 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 Res. 48 (1987) 211–221.
[15] A.M. Clark, H.J. Bohnert, Cell-specific expression of genes of the lipid transfer protein family from *Arabidopsis thaliana*, Plant Cell Physiol. 40 (1999) 69–76.
[16] G.M. Clore, M. Nilges, D.K. Sukumaran, A.T. Brnnger, M. Karplus, A.M. Gronenborn, The three-dimensional structure of α1-purothionin in solution: combined use of nuclear magnetic resonance, distance geometry and restrained molecular dynamics, EMBO J. 5 (1986) 2729–2735.
[17] T.A. Egorov, T.I. Odintsova, A. Kh. Musolyamov, R. Fido, A.S. Tatham, P.R. Shewry, Disulphide structure of a sunflower seed albumin: conserved and variant disulphide bonds in the cereal prolamin superfamily, FEBS Lett. 396 (1996) 285–288.
[18] S. Fefeu, S. Bouaziz, J.C. Huet, J.-C. Pernollet, E. Guittet, Three-dimensional solution structure of β cryptogein, a β elicitin secreted by a phytopathogenic fungus *Phytophthora cryptogea*, Protein Sci. 6 (1997) 2279–2284.
[19] J. Felsenstein, PHYLIP (Phylogeny Inference Package) Version 3.5c. Distributed by the Author, Department of Genetics, University of Washington, Seattle, 1993.
[20] C. Fernandez-Tornero, A. Ramon, M.L. Navarro, J. Varela, G. Jiménez-Gallego, Síntesis of proteins with disulfide bonds in *E. coli* using defined culture media, Biotechniques 32 (2002) 1238–1242.
[21] D.E.A. Florack, W.J. Stiekema, Thionins: properties, possible biological roles and mechanisms of action, Plant Mol. Biol. 26 (1994) 25–37.
[22] F. Garcia-Olmedo, A. Molina, A. Segura, M. Moreno, The defensive role of non-specific lipid-transfer proteins in plants, Trends Microbiol. 3 (1995) 72–74.
[23] F. Garcia-Olmedo, G. Salcedo, R. Sanchez-Monge, L. Gomez, J. Royo, P. Carbonero, Plant proteinaceous inhibitors of proteinases and α-amylases. Oxford University Press, Plant Mol. Cell. Biol. 4 (1987) 275–334.
[24] M.-F. Gautier, M.-E. Aleman, A. Guirao, D. Marion, P. Joudrier, *Triticum aestivum* puroindolines, two basic cysteine-rich seed proteins: cDNA sequence analysis and developmental gene expression, Plant Mol. Biol. 25 (1994) 43–57.
[25] M.F. Gautier, P. Cosson, A. Guirao, R. Alary, P. Joudrier, Puroindoline genes are highly conserved in diploid ancestor wheats and related species but absent in tetraploid *Triticum* species, Science 153 (2000) 81–91.
[26] M. Gijzen, S. Miller, K. Kuflu, R.I. Buzzell, B.L.A. Miki, Hydrophobic protein synthesized in the pod endocarp adheres to the seed surface, Plant Physiol. 120 (1999) 951–959.
[27] E. Gincel, J.P. Simorre, A. Caille, D. Marion, M. Ptak, F. Vovelle, Three-dimensional structure in solution of a wheat lipid-transfer protein from multidimensional ^{1}H-NMR data, Eur. J. Biochem. 226 (1994) 413–422.
[28] J. Gomar, M. Ch. Petit, P. Sodano, D. Sy, D. Marion, J.C. Kader, F. Vovelle, M. Ptak, Solution structure and lipid binding of a non-specific lipid transfer protein extracted from maize seeds, Protein Sci. 5 (1996) 565–577.
[29] R. Gonzalez, J. Varela, J. Carreira, F. Polo, Soybean hydrophobic protein and soybean hull allergy, Lancet 346 (1995) 48–49.
[30] J.H. Goo, A.R. Park, W.J. Park, O.K. Park, Selection of *Arabidopsis* genes encoding secreted and plasma membrane proteins, Plant Mol. Biol. 41 (1999) 415–423.
[31] S. Gourinath, N. Alam, A. Srinivasan, Ch. Betzel, T.P. Singh, Structure of the bifunctional inhibitor of trypsin and α-amylase from ragi seeds at 2.2 Å resolution, Acta Crystallogr. Sect. D 56 (2000) 287–293.
[32] P. Guerche, Ch. Tire, F.G. de Sa, A. De Clerq, M. Van Montagu, E. Krebbers, Differential expression of the *Arabidopsis* 2S-albumin genes and the effect of increasing gene family size, Plant Cell 2 (1990) 469–478.
[33] B. Heinemann, K.V. Andersen, P.R. Nielsen, L.M. Bech, F.M. Poulsen, Structure in solution of a four-helix lipid binding protein, Protein Sci. 5 (1996) 13–23.
[34] B. Henrissat, Y. Popineau, J.-C. Kader, Hydrophobic-cluster analysis of plant protein sequences, Biochem. J. 255 (1988) 901–905.
[35] B.J. Janssen, H.J. Schirra, F.T. Lay, M.A. Anderson, D.J. Craik, Structure of petunia hybrida defensin 1, a novel plant defensin with five disulfide bonds, Biochemistry 42 (2003) 8214–8222.
[36] C.J. Jolly, S. Rahman, A.A. Kortt, T.J.V. Higgins, Characterization of the wheat Mr 15 000 "grain softness protein" and analysis of the relationship between its accumulation in the whole seed and grain softness, Theor. Appl. Genet. 86 (1993) 589–597.
[37] M. Josè-Estanyol, P. Puigdomènech, Plant cell wall glycoproteins and their genes, Plant Physiol. Biochem. 38 (2000) 97–108.
[38] M. Josè-Estanyol, L. Ruiz-Avila, P. Puigdomènech, A maize embryo-specific gene encodes a proline-rich and hydrophobic protein, Plant Cell 4 (1992) 413–423.
[39] J.C. Kader, Lipid-transfer proteins in plants, Annu. Rev. Plant Physiol. Plant Mol. Biol. 47 (1996) 627–654.
[40] Z. Keresztessy, M.A. Hughes, Homology modelling and molecular dynamics aided analysis of ligand complexes demonstrates functional properties of lipid-transfer proteins encoded by the barley low-temperature-inducible gene family, blt4, Plant J. 14 (1998) 523–533.
[41] B. Lauga, L. Charbonnel-Campaa, D. Combes, Characterization of MZm3-3, a *Zea mays* tapetum-specific transcript, Plant Sci. 157 (2000) 65–75.
[42] F.T. Lay, F. Brugliera, M.A. Anderson, Isolation and properties of floral defensins from ornamental tobacco and petunia, Plant Physiol. 131 (2003) 1283–1293.
[43] J.Y. Lee, K. Min, H. Cha, D.H. Shin, K.Y. Hwang, S.W. Suh, Rice non-specific lipid transfer protein: the 1.6 Å crystal structure in the unliganded state reveals a small hydrophobic cavity, J. Mol. Biol. 276 (1998) 437–448.
[44] M.H. Lerche, B.B. Kragelund, L.M. Bech, F.M. Poulsen, Barley lipid-transfer protein complexed with palmitoyl CoA: the structure reveals a hydrophobic binding site that can expand to fit both large and small lipid-like ligands, Structure 5 (1997) 291–306.
[45] V. Mikes, M.-L. Milat, M. Ponchet, F. Panabières, P. Ricci, J.-P. Blein, Elicitins, proteinaceous elicitors of plant defense, are a new class of sterol carrier proteins, Biochem. Biophys. Res. Commun. 245 (1998) 133–139.
[46] L.W. Neuteboom, J.M.Y. Ng, M. Kuyper, O.R. Clijdesdale, P.J.J. Hooykaas, B.J. van der Zaal, Isolation and characterization of cDNA clones corresponding with mRNAs that accumulate during auxin-induced lateral root formation, Plant Mol. Biol. 39 (1999) 273–287.
[47] Y. Oda, T. Matsunaga, K. Fukuyama, T. Miyazaki, T. Morimoto, 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 (1997) 13503–13511.

[48] S. Odani, T. Koide, T. Ono, Y. Seto, T. Tanaka, Soybean hydrophobic protein. Isolation, partial characterization and the complete primary structure, Eur. J. Biochem. 162 (1987) 485–491.
[49] M. Oñaderra, R.I. Monsalve, J.M. Mancheño, M. Villalba, A. Martinez del Pozo, J.G. Gavilanes, R. Rodriguez, Food mustard allergen interaction with phospholipid vesicles, Eur. J. Biochem. 225 (1994) 609–615.
[50] O. Palomares, R. Monsalve, R. Rodriguez, M. Villalba, Recombinant pronapin precursor produced in *Pichia pastoris* displays structural and immunological equivalent properties to its mature product isolated from rapeseed, Eur. J. Biochem. 269 (2002) 2538–2545.
[51] D. Pantoja-Uceda, M. Bruix, G. Giménez-Gallego, M. Rico, J. Santero, Solution structure of RicC3, a 2S-albumin storage protein from *Ricinus communis*, Biochemistry 42 (2003) 13839–13847.
[52] D. Pantoja-Uceda, M. Bruix, J. Santoro, M. Rico, R. Monsalve, M. Villalba, Solution structure of allergenic 2S-albumins, Biochem. Soc. Trans. 30 (2002) 919–924.
[53] E.A. Pastorello, C. Pompei, V. Pravettoni, O. Brenna, L. Farioli, Ch. Trambaioli, A. Conti, Lipid transfer proteins and 2S-albumins as allergens, Allergy 67 (5 Suppl) (2001) 45–47.
[54] W. Paul, R. Hodge, S. Smartt, J. Draper, R. Scott, The isolation and characterisation of the tapetum-specific *Arabidopsis thaliana* A9 gene, Plant Mol. Biol. 19 (1992) 611–622.
[55] J. Poznanski, P. Sodano, S.W. Suh, J.Y. Lee, M. Ptak, F. Vovelle, Solution structure of a lipid transfer protein extracted from rice seeds. Comparison with homologous proteins, Eur. J. Biochem. 259 (1999) 692–708.
[56] K.D. Richards, R.C. Gardner, pEARLI1 (Accession No. L43080): an *Arabidopsis* member of a conserved gene family, Plant Physiol. 109 (1995) 1497.
[57] M. Richardson, Seed storage proteins: the enzyme inhibitors, in: L.J. Rogers (Ed.), Methods in Plant Biochemistry, 5, Academic Press, London, 1991, pp. 259–305.
[58] M. Rico, M. Bruix, C. Gonzalez, R.I. Monsalve, R. Rodriguez, ^{1}H NMR assignment and global fold of napin BnIb, a representative 2S-albumin seed protein, Biochemistry 35 (1996) 15672–15682.
[59] C.A. Ryan, Proteinase inhibitors in plants: genes for improving defenses against insects and pathogens, Annu. Rev. Phytopathol. 28 (1990) 425–449.
[60] R. Sanchez-Monge, L. Gomez, D. Barber, C. Lopez-Otin, A. Armentia, G. Salcedo, Wheat and barley allergens associated with baker's asthma. Glycosylated subunits of the α-amylase-inhibitor family have enhanced IgE-binding capacity, Biochem. J. 281 (1992) 401–405.
[61] P.R. Shewry, J.A. Napier, A.S. Tatham, Seed storage proteins: structures and biosynthesis, Plant Cell 7 (1995) 945–956.
[62] P.R. Shewry, A.S. Tatham, The prolamin storage proteins of cereal seeds: structure and evolution, Biochem. J. 267 (1990) 1–12.
[63] D.H. Shin, J.Y. Lee, K.Y. Hwang, K.K. Kim, S.W. Suh, High-resolution crystal structure of the non-specific lipid-transfer protein from maize seedlings, Structure 3 (1995) 189–199.
[64] B.S. Shorrosh, L. Wen, K. Ch. Zen, J.-K. Huang, J.S. Pan, M.A. Hermodson, K. Tanaka, S. Muthukrishnan, G.R. Reeck, A novel cereal storage protein: molecular genetics of the 19 kDa globulin of rice, Plant Mol. Biol. 18 (1992) 151–154.
[65] P. Sodano, A. Caille, D. Sy, G. de Person, D. Marion, M. Ptak, ^{1}H NMR and fluorescence studies of the complexation of DMPG by wheat non-specific lipid transfer protein. Global fold of the complex, FEBS Lett. 416 (1997) 130–134.
[66] P. Sterk, H. Booij, G.A. Schellekens, A. Van Kammen, S.C. De Vries, Cell-specific expression of the carrot EP2 lipid transfer protein gene, Plant Cell 3 (1991) 907–921.
[67] S. Strobl, K. Maskos, G. Wiegand, R. Huber, F.X. Gomis-Rüth, R. Glockshuber, 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 (1998) 911–921.
[68] R.M. Sweet, H.T. Wright, J. Janin, C.H. Chotia, D.M. Blow, Crystal structure of the complex of porcine trypsin soybean trypsin inhibitor (Kunitz) at 2.6 Å resolution, Biochemistry 13 (1974) 4212–4228.
[69] S. Tassin, W.F. Broekaert, D. Marion, D.P. Acland, M. Ptack, F. Vovelle, P. Sodano, Solution structure of Ace-AMP1, a potent antimicrobial protein extracted from onion seeds. Structural analogies with plant non-specific lipid transfer proteins, Biochemistry 37 (1998) 3623–3637.
[70] A.S. Tatham, P. Masson, Y. Popineau, Conformational studies of peptides derived by the enzymic hydrolysis of a gamma-type gliadin, J. Cereal Sci. 11 (1990) 1–13.
[71] F.R.G. Terras, I.J. Goderis, F. Van Leuven, J. Vanderleyden, B.P.A. Cammue, W.F. Broekaert, In vitro antifungal activity of a radish (*Raphanus sativus* L.) seed protein homologous to nonspecific lipid transfer proteins, Plant Physiol. 100 (1992) 1055–1058.
[72] S. Thoma, U. Hecht, A. Kippers, J. Botella, S. De Vries, C. Somerville, Tissue-specific expression of a gene encoding a cell wall-localized lipid transfer protein from *Arabidopsis*, Plant Physiol. 105 (1994) 35–45.
[73] S. Thoma, Y. Kaneko, C. Somerville, A non-specific lipid transfer protein from *Arabidopsis* is a cell wall protein, Plant J. 3 (1993) 427–436.
[74] S. Udenfriend, K. Kodukula, How glycosyl-phosphatidy linositol-anchored membrane proteins are made, Annu. Rev. Biochem. 64 (1995) 563–591.
[75] F. Vallée, A. Kadziola, Y. Bourne, M. Juy, K.W. Rodenburg, B. Svensson, R. Haser, Barley α-amylase bound to its endogenous protein inhibitor BASI: crystal structure of the complex at 1.9 Å resolution, Structure 6 (1998) 649–659.
[76] H. Van der Klei, J. Van Damme, P. Casteels, E. Krebbers, A fifth 2S-albumin isoform is present in *Arabidopsis thaliana*, Plant Physiol. 101 (1993) 1415–1416.
[77] F. Vignols, M. Wigger, J.M. Garcia Garrido, F. Grellet, J.C. Kader, M. Delseny, Rice lipid transfer protein (LTP) genes belong to a complex multigene family and are differentially regulated, Gene 195 (1997) 177–186.
[78] R.H. Voss, U. Ermler, L.-O. Essen, G. Wenzl, Y.-M. Kim, P. Flecker, Crystal structure of the bifunctional soybean Bowman–Birk inhibitor at 0.28-nm resolution. Structural peculiarities in a folded protein conformation, Eur. J. Biochem. 242 (1996) 122–131.
[79] Y.-M. Woo, D.W.-N. Hu, B.A. Larkins, R. Jung, Genomics analysis of genes expressed in maize endosperm identifies novel seed proteins and clarifies patterns of zein gene expression, Plant Cell 13 (2001) 2297–2317.