CMLS, Cell. Mol. Life Sci. 58 (2001) 556–570
1420-682X/01/040556-15 $ 1.50 + 0.20/0

CMLS Cellular and Molecular Life Sciences

Review

Revisiting the enzymes stored in the laticifers of *Carica papaya* in the context of their possible participation in the plant defence mechanism

A. El Moussaoui, M. Nijs, C. Paul, R. Wintjens‡, J. Vincentelli, M. Azarkan and Yvan Looze*

Université Libre de Bruxelles, Campus Erasme, CP 609, 808, Route de Lennik, B-1070, Bruxelles (Belgium), Fax + 32 2 555 6782, e-mail: ylooze@ulb.ac.be

Received 8 May 2000; received after revision 25 September 2000; accepted 25 September 2000

Abstract. In the tropical species *Carica papaya*, the articulated and anastomosing laticifers form a dense network of vessels displayed in all aerial parts of the plant. Damaging the papaya tree inevitably severs its laticifers, eliciting an abrupt release of latex. Besides the well-known cysteine proteinases, papain, chymopapain, caricain and glycyl endopeptidase, papaya latex is also a rich source of other enzymes. Together, these enzymes could provide an important contribution to plant defence mechanisms by sanitising and sealing the wounded areas on the tree.

Key words. Laticifer; *Carica papaya*; papain; 5-oxoproline; Kunitz; lipase; glycosyl hydrolase; phytocystatins.

Introduction

Plants must fend off a myriad of aggressive biotic agents. This biotic onslaught occurs across the complete spectrum of spatial scales from pathogens that attack single cells to insect herbivores that select their hosts by characteristics of the community in which a plant grows [1, 2]. Several thousands of plant species make use of specialised structures called *laticifers* to secure, at least partly, their defence.

There are two main types of laticifers. Non-articulated laticifers initiate as single cells that push their way between other cells while growing. They continue to develop as the plant grows, to become giant coenocytic cells. Articulated laticifers, on the other hand, begin as a column of cells. During development, they form vessel-like structures through breakdown of adjoining cell walls. In many species, neighbouring vessels anastomose. Both types of laticifer form long and narrow structures.

In many laticiferous plants, latex is stored under pressure, which implies that, in response to cuts and/or bites by plant predators, it is abruptly expelled. Experimental evidence, accumulated in the course of several decades, has shown that latex contributes to protecting the plant against predators. Such a function can be successfully achieved either mechanically or chemically. For example, as a result of coagulation, latex droplets can eventually harden to the point of virtually muzzling predators [3]. In addition, toxic substances, when present in latex, have been shown to have a negative impact on both insect feeding rate and predator fitness [4].

Laticifers present an intensive metabolic activity. In situ, they contain the usual organelles of plant cells including nuclei, mitochondria, plastids, vacuoles, ribosomes, Golgi apparatus and endoplasmic reticulum [5]. Botanists designate the cytoplasm of laticiferous structures as latex. From a chemical point of view, it is a complex mixture of various chemical substances.

* Corresponding author.
‡ Present address: Institut Pasteur de Lille, CNRS UMR-8525, Lille (France).

In common with other types of plant cell, laticifers can synthesise the numerous molecules required to achieve their basic physiological functions. However, they distinguish themselves by their ability to synthesise and store secondary metabolites, often present in appreciable amounts. These metabolites display chemical structures that greatly differ according to the family to which the plant species belongs. The latex of the poppy plant (*Papaver somniferum*; family: Papaveraceae), for example, is well known as a rich source of some 20 alkaloids, among which morphine and codeine are acknowledged as powerful analgesics. In contrast, laticifers of the rubber tree (*Hevea brasiliensis*; family: Euphorbiaceae) are equipped for the efficient synthesis and storage of *cis*-1,4-polyisoprene that may represent up to 30% of dried *Hevea* latex.

Plant species belonging to several families (Apocynaceae, Asteraceae, Caricaceae, Euphorbiaceae and Moraceae) have been observed to take advantage of their laticifers to fend off predators [3]. How the secondary metabolites present in their laticifers can fulfil a universal function while exhibiting such different chemical structures is intriguing. This question is addressed here for the case of the tropical species *Carica papaya* (family: Caricaceae). This case is perplexing, because the generally accepted opinion is that proteinase inhibitors (PIs) actively contribute to plant defence mechanisms [2]. Contradicting this view, *C. papaya* and several other *Carica* species store tremendous amounts of cysteine proteinases (EC 3.4.22) in their laticifers [6, 7].

C. papaya is a soft-stemmed and unbranched tree able to grow to 20 m in height. Its large leaves, by emerging directly from the upper part of the stem on long petioles, confer a palm-like appearance to the tree. Being soft-stemmed, this species, at first sight, also appears as somewhat flimsy and thus a facile prey. Nonetheless, native to Central America, the papaya tree has successfully established in many ecological niches. It is now widely cultivated in tropical and subtropical regions around the world for its edible fruit and its latex (see fig. 1). This successful acclimatisation is a strong indication that the papaya tree is able to compensate efficiently for the absence of lignin.

Laticifers are displayed in all aerial parts of the tree forming a dense network of *articulated* and *anastomosing* structures [8]. Such a layout is particularly efficient, since it prevents the latex flow, within the laticifers, from being interrupted. In *Carica* leaves, for example, damaging the venation network is of limited consequence because latex can bypass severed veins by flowing through neighbouring ones.

Papaya latex is a thixotropic fluid with a milky appearance that contains about 85% water. An insoluble particulate fraction whose composition is still practically unknown, makes up 25% of the dry matter. The soluble fraction, on the other hand, contains both the usual ingredients such as carbohydrates ($\sim 10\%$), salts ($\sim 10\%$) and lipids ($\sim 5\%$), and representative biomolecules such as glutathione, cysteine proteinases ($\sim 30\%$) and several other proteins ($\sim 10\%$).

The papaya proteinases have been widely used for several decades in the food (e.g. for meat tenderisation and beer chill-proofing) and pharmaceutical industries. Hence, the economic interest in papaya latex has prompted much research aimed at elucidating both the structures and the modes of action of the papaya proteinases.

Although papain [EC 3.4.22.2] is a minor constituent among the papaya proteinases, this enzyme has been the most extensively studied. Its primary structure was fully elucidated as in 1970 [9] and it was the second enzyme to be crystallised and have its structure determined by X-ray crystallography [10]. Joe R. Kimmel,

Figure 1. *Carica papaya* is cultivated for its edible fruit and latex, a rich source of proteinases used in the food and pharmaceutical industries. Fully grown but unripe fruits constitute the source of latex that is produced at the rate of several thousand metric tons each year. An orchard of not less than 10 ha is normally required to produce one metric ton of dry latex annually.

Table 1. Characterised proteins and enzymes present in papaya laticifers.

	Isoform	Accession number		
		Swiss Prot	EMBL (cDNA)	PDB (latest entry)
Papain		P00784	M15203	9PAP
Chymopapain	1	P14080	X97789	1YAL
	2		AJ131995	
	3		AJ131996	
	4		AJ131997 (p)	
	5		AJ131998 (p)	
Caricain	1	P10056	X66060	1PPO
	2		not deposited	
Glycyl endopeptidase		P05994	X78056	1GEC
Cystatin			X71124	
Kunitz-type inhibitor		P80691		
β-1,3-Glucanase				
Lysozyme		P81241 (p)		
Class II chitinase		P81241 (p)		
Glutaminyl cyclase			AF061240	
Lipase (s)				

(p) only partial sequences are available.

Emil L. Smith and Jan Drenth have associated their names with these pioneering studies on papain.

Full characterisation of the other papaya proteinases has long been hampered by difficulties encountered in the course of their purification [11–13]. The revival in papaya latex enzyme research that took place during the last decade had to await the discovery of papaya proteinase IV (also known as glycyl endopeptidase and chymopapain M) [14] and the use of molecular biology techniques. Much progress in the knowledge of papaya proteinases has accumulated during this period, considerably increasing the number of known papaya latex enzymes that are now fairly well characterised. An exhaustive list of the papaya enzymes that are reviewed here is shown in table 1.

The papaya glycosyl hydrolases

Glycosyl hydrolases are a diverse group of enzymes with respect to their enzymatic activities and hydrolytic mechanisms. They may exert important functions in plant metabolism but additionally participate in plant responses to invasion by pathogenic micro-organisms [15].

Several glycosyl hydrolases have been identified in papaya latex. A β-1,3-endoglucanase (EC 3.2.1.39, family 17) has been characterised as a basic protein that exhibits a marked preference for substrates containing long sequences of adjacent β-1,3-linkages [16]. This enzyme has only been partially purified due to its high susceptibility to proteinases [16].

The first chitinase was purified to homogeneity and characterised, three to four decades ago, as an enzyme capable of hydrolysing not only chitin but also peptidoglycans from bacterial cell walls [17–22]. This enzyme, by cleaving peptidoglycans between the C-1 of a muramic acid residue and the C-4 of N-acetylglucosamine [19], exhibits the characteristics of a true lysozyme. On such a substrate, cleavage by the papaya enzyme proceeds with inversion of anomeric configuration [20]. From a physicochemical point of view, papaya lysozyme is a basic protein with a molecular mass of 28 kDa [17–19]. Its polypeptide chain is stabilised by two to three disulphide bonds and also contains several accessible free thiol functions [21, 22].

Glycosyl hydrolases, on the basis of sequence similarities, have been classified in a number of families with well-defined enzymatic properties [23]. Chitinases belong either to family 18 or 19 of glycosyl hydrolases. Another classification of plant chitinases discerns five classes, of which classes I, II and IV have homologous sequences and constitute family 19 of the glycosyl hydrolases, while representatives of classes III and V belong to glycosyl hydrolases family 18 [24]. Chitinases belonging to classes I and IV have a separate N-terminal chitin-binding domain, linked to the catalytic domain by a hinge peptide sequence. By definition, class II chitinases lack the hevein (chitin-binding) domain [24].

Using a combination of ion exchange and hydrophobic interaction chromatographies, a 28-kDa chitinase has recently been purified from papaya latex and unambiguously classified as a class II enzyme [25]. Identification as a class II chitinase was mainly based on the knowledge of its N-terminal amino acid sequence determined up to the 30th amino acid residue by automatic Edman degradation. Of note, this papaya chitinase did not

possess any measurable bacteriolytic activity against *Micrococcus luteus* cells, an observation consistent with the fact that (to the best of our knowledge) no class II chitinase has yet been reported to possess bacteriolytic activity. This and several other observations suggest that papaya class II chitinase and the formerly identified papaya lysozyme are distinct molecular species. The three free thiol functions in papaya chitinase, in marked contrast to those in papaya lysozyme, are completely buried and only become amenable to chemical modification after the polypeptide chain has been unfolded [22, 25]. Furthermore, both enzymes show somewhat different affinities towards several chromatographic supports [25; unpublished observations]. On the other hand, crystal structures for a class II chitinase (CHI-2) from *Hordeum vulgare* and a class III chitinase (hevamine) from *H. brasiliensis* have both been refined at 1.8-Å resolution [26, 27]. Interestingly, these chitinases adopt quite different folds leading to classification of CHI-2 as an all-α and hevamine as a $(\beta\alpha)$8-barrel protein (fig. 2). Obviously, the circular dichroism spectrum of papaya lysozyme, as reported by Barel et al. [21], is not consistent with those displayed by all-α proteins which also argues against the classification of this papaya enzyme within class II chitinases. A preparation of class II chitinase with identical N-terminal sequence [25] was found to have lysozyme activity [28]. These authors suggested that papaya lysozyme is identical to the class II chitinase, also because the reported five N-terminal residues differ at only one position [19, 25, 28]. But this is not in agreement with other substantial evidence as summarised above. Bacteriolytic activity has often been found associated with class I [15] or class III chitinases, in, for example, lys b1, b2 and 28a from *N. tabacum* [15] and hevamine itself [29]. The N-terminal amino acid sequence, reported for papaya lysozyme [19], extends only to the fifth amino acid residue which, however, provides sufficient information to exclude the possibility that papaya lysozyme should belong to class I chitinases.

Unlike papaya lysozyme, however, hevamine is not a true lysozyme since it cleaves proteoglycan between C-1 of N-acetylglucosamine and C-4 of N-acetylmuramic acid [29]. Additionally, unlike the papaya enzyme and class I chitinases, cleavage of chito-oligosaccharides by the class III chitinases from cucumber leaves and the *Hevea* tree proceeds with retention of anomeric configuration [30, 31]. Such dissimilarities, possibly explained by subtle differences in the active site pockets, should, therefore, preclude classifying papaya lysozyme within the class III chitinases.

Altogether, the papaya glycosyl hydrolases, especially in combination, can inhibit the growth of several classes of micro-organisms. The substrates of β-1,3-glucanases and chitinases, β-1,3-glucan and chitin (a polymer of N-acetylglucosamine), respectively, are present in the cell walls of many fungi. Peptidoglycans, the substrates of lysozyme, are the main constituents of bacterial walls. A first contribution by these enzymes to plant defence thus results from their ability to degrade the cell walls of invading fungal [32] and bacterial pathogens.

β-Glucan and chitin-like oligosaccharides are potent inducers of several other plant defence reactions [15]. Hence, a second potential function of β-1,3-glucanase and chitinase could also involve release of defence-triggering signal molecules (the so-called elicitors).

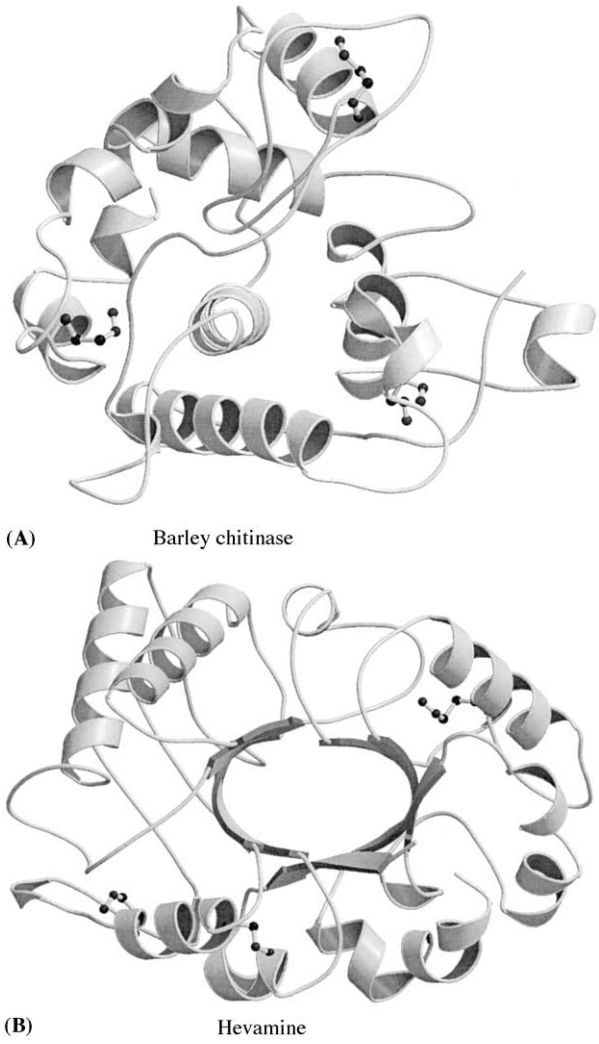


Figure 2. Ribbon drawings [98, 99] of the X-ray structures of barley chitinase [26] (*A*) and hevamine [27] (*B*). Disulphide bonds are depicted in 'ball-and-stick' style.

Table 2. Sequence homology within the family of phytocystatins.

	Carica papaya	*Oryza sativa* I	*Oryza sativa* II	*Zea mays* I*	*Zea mays* II*	*Vigna unguiculata*	*Glycine max* I*	*Glycine max* II	*Ambrosia artemisiifoli*	*Castanea sativa*	*Persea americana*	*Helianthus annuus*	*Daucus carota**	*Chelidonium majus*
Papaya (EMBL X71124)		*90*	*85*	*90*	*90*	*92*	*95*	*90*	*90*	*93*	*93*	*75*	*97*	*78*
Rice I (SP 09229)	**49**		*86*	*103*	*103*	*94*	*88*	*92*	*90*	*88*	*95*	*81*	*90*	*89*
Rice II (SP 20907)	**52**	**65**		*89*	*85*	*91*	*88*	*91*	*78*	*85*	*84*	*78*	*55*	*73*
Maize I* (SP 31726)	**51**	**67**	**64**		*135*	*88*	*116*	*87*	*90*	*88*	*95*	*81*	*131*	*88*
Maize II* (EMBL D38130)	**53**	**71**	**62**	**87**		*89*	*115*	*87*	*91*	*88*	*95*	*81*	*124*	*89*
Cowpea (SP Q06445)	**59**	**68**	**57**	**64**	**64**		*97*	*92*	*92*	*97*	*92*	*80*	*98*	*76*
Soybean I* (EMBL D64115)	**50**	**61**	**55**	**51**	**52**	**68**		*92*	*90*	*100*	*99*	*76*	*100*	*90*
Ragweed (EMBL L16624)	**48**	**54**	**53**	**56**	**56**	**61**	**52**	**62**		*91*	*90*	*77*	*94*	*78*
Chestnut tree (EMBL AJ224331)	**59**	**64**	**55**	**65**	**67**	**73**	**59**	**74**	**53**		*98*	*82*	*107*	*76*
Avocado (PIR JHO269)	**55**	**61**	**60**	**64**	**65**	**68**	**69**	**66**	**62**	**61**		*82*	*99*	*89*
Sunflower (SP Q10992)	**61**	**65**	**61**	**68**	**70**	**75**	**73**	**73**	**70**	**68**	**72**		*85*	*75*
Carrot* (EMBL D85623)	**31**	**39**	**39**	**35**	**32**	**33**	**29**	**32**	**33**	**35**	**35**	**39**		*101*
Celandine [28]	**39**	**40**	**43**	**43**	**44**	**34**	**40**	**38**	**42**	**33**	**49**	**51**	**60**	

The number of amino acid residues that were aligned in both compared sequences is shown in italic type while percent identity is shown in bold type. Asterisks denote deduced sequences that contain a putative signal sequence.

The papaya protease inhibitors

It is now well established that plants make use of PIs as defensive proteins against predators. After attack, both long- (systemin) and short- (cell wall fragments) distance cascades result in transcriptional up-regulation of PI gene expression. Furthermore, plants genetically engineered with PI-encoding cDNA sequences are better protected against natural enemies [2]. PIs are generally thought to function as digestibility reducers by targeting the major proteolytic digestive enzymes of herbivores.

Two such PIs are stored in the laticifers of *C. papaya*. A papaya cystatin-encoding cDNA clone has been isolated from a papaya leaf cDNA library and sequenced. The predicted amino acid sequence revealed a single polypeptide chain that contains 98 amino acid residues (Mr = 11,131 Da). Its encoding cDNA clone was isolated as a result of screening the leaf cDNA library with polyclonal rabbit antibodies raised against proteins derived from papaya latex [33]. Papaya cystatin is therefore expected to be stored in papaya laticifers even though it has not yet been isolated from papaya latex. The absence of a signal peptide and of putative glycosylation site(s) in the predicted sequence [33] strongly suggests that this inhibitor is located in the cytoplasm of laticifer cells where it could possibly prevent undesirable vessel damage by endogenous proteinases.

Papaya cystatin exhibits amino acid sequence homology with cysteine PIs, especially those belonging to the phytocystatins group (table 2). Most members from this group, including papaya cystatin (theoretical pI 5.33) are acidic proteins, although some of them, such as chelidocystatin (pI 9.3), are basic [34]. Phytocystatins with and without putative signal peptides have been identified. This strongly suggests that they may fulfil distinct physiological functions as a result of their delivery to different locations. The phytocystatin group thus does not appear as a homogeneous collection. Moreover, members of this group exhibit sequence homology with animal cystatins that belong to types 1 and 2. The amino acid sequence of papaya cystatin, for example, displays 33 and 28% identity with hen egg white cystatin (type 2; Swiss Prot access number P01038) and human stefin B (type 1; Swiss Prot access number P04080), respectively. However, phytocystatins have been gathered together into a subfamily less on the basis of sequence homologies, but more because their gene organisation is generally different from that of animal cystatins [35]. Furthermore, despite their closer similarity with animal cystatin type 2 members, phytocystatins constitute a subgroup in type 1 of animal cystatins because, like stefin B, they lack disulphide bonds and have shorter amino acid sequences.

Papaya cystatin also inhibits papain (Kd = 0.39 nM), chymopapain (Kd = 0.43 nM), caricain (Kd = 1.49 nM)

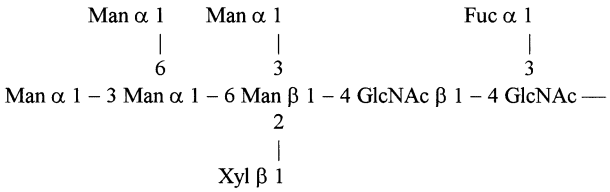
and glycyl endopeptidase (Kd = 3.44 nM) [33]. Its affinity towards papain is quite comparable to that displayed by several other members of the phytocystatin subfamily. Cystatins from *Chelidonium majus* and *Malus domestica*, for example, bind to this protein with Kds of 0.11 and 0.21 nM, respectively [34, 36]. Papain is, however, still much more efficiently inhibited by animal cystatins. Kd values of 19 pM and even of 60 fM have been reported for human stefin A and chicken cystatin, respectively [37, 38].

Papaya cystatin does not discriminate strongly between the four papaya proteinases, contrasting with the behaviour exhibited by all other cystatins examined so far. Glycyl endopeptidase, which forms a tight complex with papaya cystatin, is not inhibited at all in the presence of chicken cystatin [14] nor in the presence of the phytocystatin purified from potato tubers [39]. Ratios of Kd (chymopapain) to Kd (papain) of 12.5, 64 and >1000 have been reported for the cystatins from chestnut seeds [40], potato tubers and hen egg white [39], respectively, even though both papaya proteinases are not discriminated at all by the papaya inhibitor. Similarly, the ratio of Kd (caricain) to Kd (papain) is only 4 in the case of papaya cystatin against 45.5 and $>10{,}000$ for potato and chicken cystatins, respectively [39].

Another unique structural feature exhibited by papaya cystatin concerns the conserved motif Gln-Xaa-Val-Xaa-Gly present in about the middle of the cystatin sequences (segment 52–56 in the amino acid sequence of papaya cystatin). The motif in animal [41, 42] and plant cystatins [43] contributes to interactions with papain. The strictly conserved glutamine residue in this motif is substituted by alanine in the papaya protein. In the refined crystal structure of recombinant human stefin B in complex with papain, the side chain of the generally conserved glutamine residue makes a polar contact with the carbonyl oxygen atom of residue 63 in the papain amino acid sequence [42]. Given that tight binding of cystatins to papain involves many such contacts [42, 44], substituting glutamine 52 by alanine is expected to have a negligible effect on the inhibitory capacity of papaya cystatin.

Affinity chromatography of papaya latex on immobilised trypsin provides a nearly homogeneous ($>95\%$) preparation of a polypeptide chain (called here PPI) exhibiting the ability to inhibit both bovine trypsin and chymotrypsin. Its complete primary structure, determined by protein analysis, revealed that PPI belongs to the superfamily of the Kunitz-type trypsin inhibitors. The covalent structure of PPI consists of 184 amino acid residues including four cysteinyls paired into two disulphide bonds. In the course of sequencing work, the PPI preparation was demonstrated to consist of an equimolar mixture of two polypeptide chains that differed solely in the presence or absence of an N-terminal valyl residue [45].

PPI is a glycoprotein as predicted by the existence of two putative N-glycosylation sites located at positions 86 and 92 in the amino acid sequence. The structure of the major glycan moiety that was released from the protein by hydrazinolysis, with its five mannose residues, as shown below, represents a new plant-type carbohydrate chain [46].



Experimental evidence strongly suggests that both putative N-glycosylation sites are actually glycosylated, despite their close proximity within the amino acid sequence.

To visualise the positions occupied by these glycans at the three-dimensional (3D) level, both the 3D structure of the

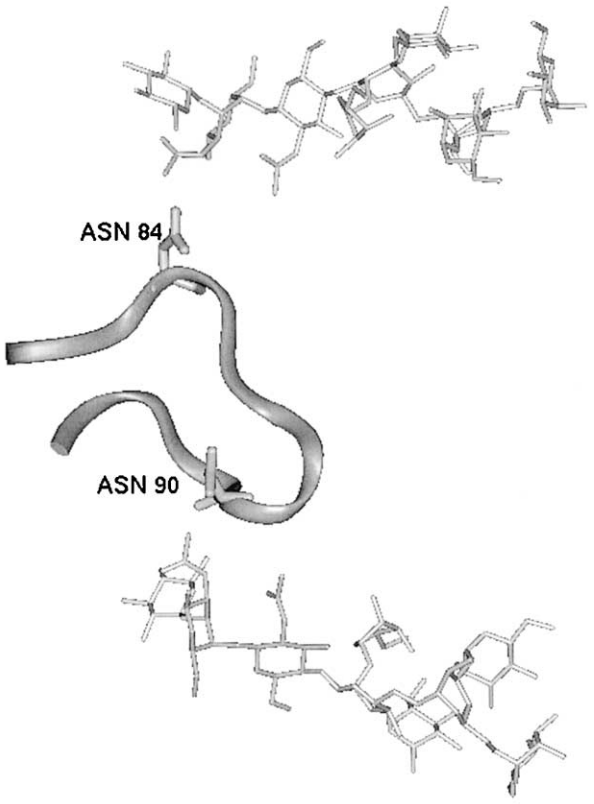


Figure 3. Close-up view of the glycosylated loop of PPI. The loop is shown as a ribbon and the two N-glycosylation sites (Asn 84 and Asn 90) are detailed and labelled. The 3D model of PPI was built by homology modelling [100] based on the X-ray crystal structures of trypsin inhibitors from soybean [101], *E. caffra* [102] and winged bean [103, 104] seeds. The carbohydrate chains [46] are depicted in 'ball-and-stick' style.

major carbohydrate chain and that of PPI itself have been modelled. Full details of this investigation will be reported elsewhere. Briefly, the oligosaccharide was constructed iteratively by stepwise addition of monosaccharide units starting from the first GlcNAc residue and using the molecular modelling package Insight II. Prediction of the 3D structure of PPI, using the automated program MODULAR, was based on satisfaction of spatial restraints derived from structures of homologous proteins. The radiocrystallographic structures of the Kunitz-type PIs from *Glycine max*, *Erythrina caffra* and *Psophocarpus tetragonolobus* (PDB ID codes: 1AVU, 1TIE, 1WBA and 1WBC) were used for this purpose. According to the predicted structure, both glycosylated sites are located on a long external loop constituted by the PPI fragment starting with Asn 86 and ending with Ala 97. As shown in figure 3, both oligosaccharides, because they move towards different directions, could be constructed without any overlap.

The carbohydrate chains, spread out at the interface between the surface of the PPI molecule and the bulk of the solvent, reduce the solvent accessibility of the protein by about 1400 Å^2 (about 15% of the total). As a result, loop 86–97 in PPI is well protected, for example, against proteolytic attack. The glycans, however, are also able to influence more distant regions, such as neighbouring loops, in the PPI molecule. One such loop is the reactive site loop that starts with Asn 67 and ends at Leu 79 and contains Lys 69 as the sound P1 residue. The presence of carbohydrate chains in close proximity might, thus, also explain why PPI is only moderately active against bovine trypsin (Kd = 300 nM) and bovine chymotrypsin (Kd = 800 nM) and also why peptide bonds in the reactive site loop have so far resisted hydrolysis by both serine proteinases [45].

Overall, the 3D model of PPI consists of 12 antiparallel β strands connected by long loops, the β strands forming the buried core of the protein, and the loop regions, the solvent-accessible surface of the molecule. Hence, PPI may be regarded as an all-β protein with 33.2% of its residues in β strand conformation (fig. 4). The predicted structure of papaya cystatin (also shown in fig. 4), based on that of human stefin B [42], reveals quite different folding. Furthermore, inhibition of proteases by papaya cystatin and PPI imply different modes of interaction. As already mentioned, one single loop (residues 67–79 in PPI) is implicated in the inhibition of serine proteinases [45]. In marked contrast, several regions, distant in the amino acid sequences of cystatins, participate in the interaction with cysteine proteinases [41, 42].

Production of PIs following herbivore attack appears to be a universal plant response. However, many herbivores are able to thwart such a strategy by adopting compensatory responses, e.g. by altering the composi-

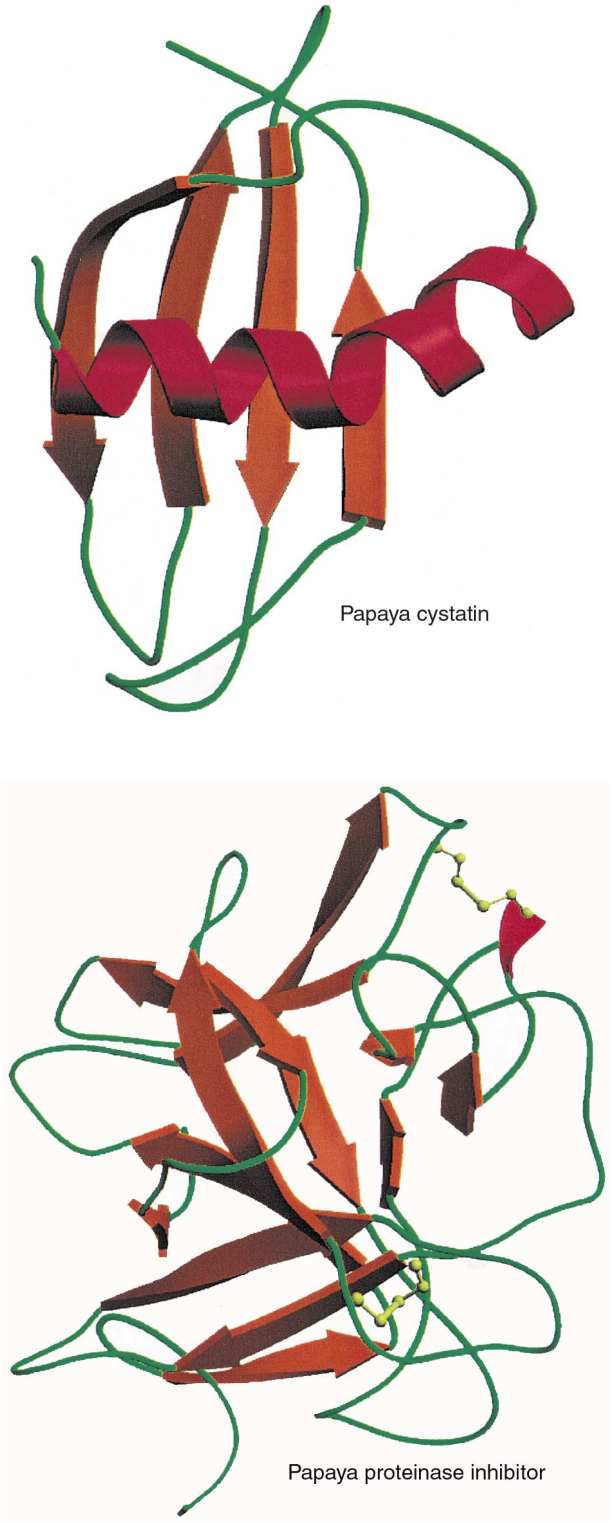


Figure 4. Ribbon models of papaya cystatin (*A*) and PPI (*B*), both predicted by homology modelling [100]. The structure of human stefin B [42] was exploited to build the 3D structure of papaya cystatin. α helices are represented as red ribbons, β strands as orange arrows. Coil regions are in green and disulphide bonds are detailed in yellow. Generated using a combination of Molscript [98] and Raster 3D [99] programs.

tion of the proteolytic enzymes they secrete into their digestive cavity [2]. This raises the question of the long-term usefulness of protease inhibition in plant protection. More recent findings nevertheless suggest that PIs can contribute to plant defence mechanisms thanks to their antiviral and antibacterial activities [47].

The papaya cysteine proteinases

So far, four proteinases, namely papain (EC 3.4.22.2), chymopapain (EC 3.4.22.6), caricain (EC 3.4.22.30) (formerly known as proteinase omega) and glycyl endopeptidase or papaya proteinase IV (EC 3.4.22.25) have been isolated from the latex of *C. papaya*. They all belong to the peptidase family C1, also known as the papain clan of cysteine proteinases. Their amino acid sequences have been determined both at the protein level [9, 48–50] and through sequencing corresponding cDNA clones [51–54].

Papaya proteinases are all synthesised as proenzymes. The mature forms of the proteinases generally contain from 212 to 218 amino acids and are preceded by 106- to 108-residue long prosequences (fig. 5). The four sequences exhibit a strong degree of homology. Identification of putative N-glycosylation sites on three of these prosequences provides evidence that, after synthesis, they possibly transit through the Golgi apparatus. In any case, polyribosomes attached to membranes of the endoplasmic reticulum do appear to be the site of their biosynthesis as testified by the presence of signal peptides (not shown in fig. 5) containing from 16 to 18 amino acid residues. Two similar but distinct cDNAs have been shown to code for caricain [52]. The deduced amino acid sequence shown in figure 5 is identical to the experimentally determined amino acid sequence of caricain [48]. The other differs at only one position (T 214 I) and, distinctively, contains a 19-amino-acid-long C-terminal extension. Furthermore, at least five similar but distinct cDNAs code for chymopapain [54], among

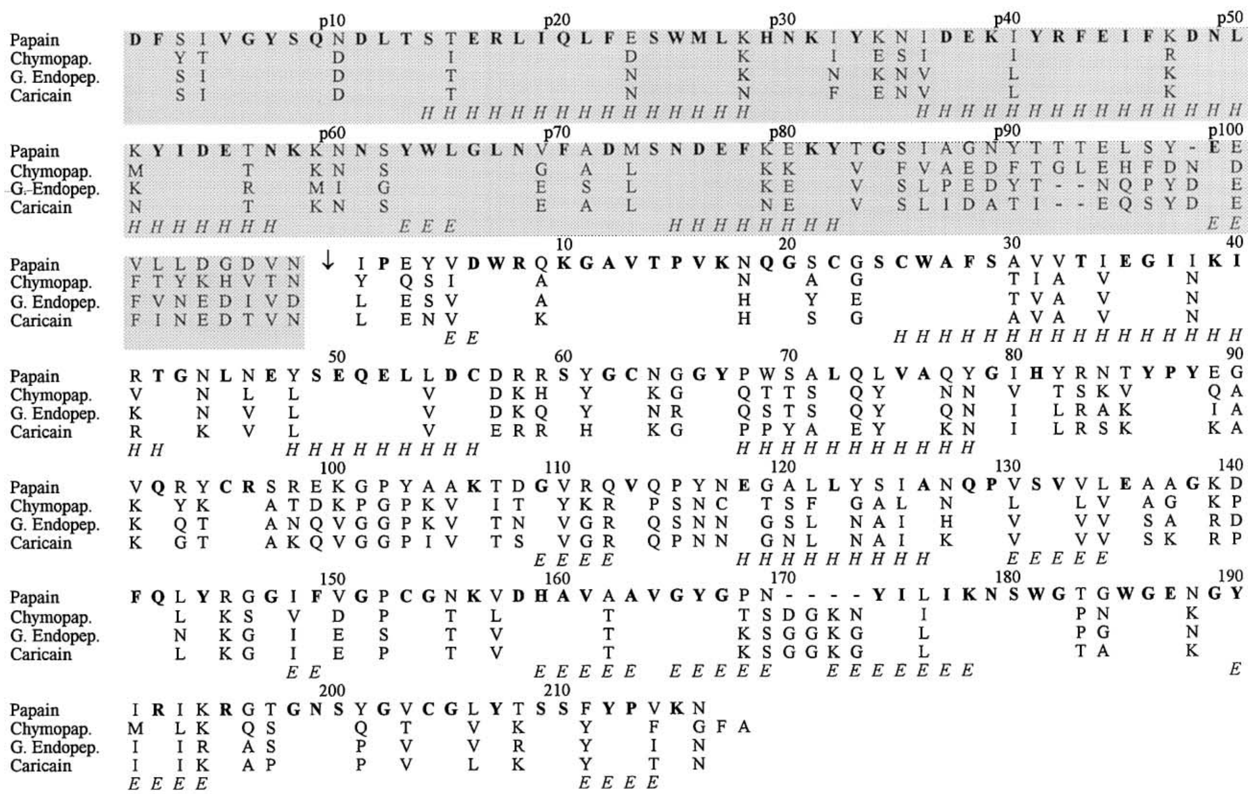


Figure 5. Alignment of the prosequences of *C. papaya* cysteine proteinases. The protein sequences of papain, chymopapain (isoform 1, Chymopap.), glycyl endopeptidase (G. Endopep.) and caricain (isoform 1) are shown. Residues conserved in all four sequences are in bold and, for clarity, only typed in the propapain sequence; deletions are denoted by hyphens. The arrow indicates the cleavage site delimiting the proregion (shaded) from the mature protein. The secondary-structure elements identified in the proregion of caricain [73] as well as those conserved in the radiocrystallographic structures of the mature enzymes [56–59] are also shown: residues that participate in β ladders and those that contribute to α helices are indicated by the italised letters E and H, respectively. References to original sequencing work are given in the text.

which three display C-terminal extensions containing 9 amino acid residues missing in the sequenced protein shown in figure 5 [49, 55].

Comparison of the X-ray structures of the four papaya proteinases reveals that they adopt identical 3D folds [56–59]. A close examination of their active site grooves also shows that the specificity restrictions dictated by the S2 subsites do not differ significantly among the four proteinases. However, contrasting with the situation in papain, chymopapain and caricain, the active site of papaya proteinase IV is occluded due to the substitution of the highly conserved residues Gly 23 and Gly 65 by Glu and Arg, respectively, [50]. As a result, papaya proteinase IV is a highly specific endoproteinase for the peptide linkage Gly-Xaa [60], while papain, chymopapain and caricain are rather general endopeptidases [61].

In papaya laticifers, the four cysteine proteinases together account for about 80% of the whole enzyme fraction. Assuming they can circulate freely within the laticifers, their concentration has been calculated to be higher than 1 mM [62]. The papaya proteinases thus constitute a potential danger for the plant, leading to the suggestion that they must necessarily be sequestered within some organelles [52]. Close examination of their chemical structures, however, has not clearly identified sorting signals that would precisely indicate which organelle is the actual site of storage. On the other hand, active proteinases, in freshly collected papaya latex, only represent a fraction of the total potential proteolytic activity [63], and chymopapain has been tentatively identified as the papaya proteinase active in vivo [64]. These observations by Salas and co-workers strongly suggest that (i) a substantial fraction of the pool of papaya proteinases, in vivo, is stored somewhere in an inactive form but also that (ii) the intracellular trafficking of these proteinases may well differ among molecular species.

Although proforms of proteinases have not been isolated from papaya latex, they have been successfully expressed in heterologous hosts. The proform of papain has been expressed in insect cells [65], *Escherichia coli* [66] and *Saccharomyces cerevisiae* [67]. The proforms of the other three proteinases have also been produced by *E. coli* cells [52–54].

In all but one case (glycyl endopeptidase), heterologous cells express inactive precursors. *E. coli* invariably produces recombinant proproteinases entrapped in protein bodies from which they can be dissolved in the presence of chaotropic agents and subsequently refolded. In vitro conversion of the folded precursors into fully active proteinases can then proceed autocatalytically at pH 4 [52, 54, 66]. Activation of propapain appears to require several cleavage steps. One of the possible mechanisms suggested by the work of Vernet et al. [68] involves the initial cleavage of the proregion of papain at position p72 (see fig. 5) followed by a second cleavage which removes the remnant of the proregion. Full details of the stages involved in the maturation process, however, still remain obscure. Unlike the above-mentioned precursors, propapaya proteinase IV is active against synthetic substrates such as Boc-Ala-Ala-Gly-p-nitroanilide while exhibiting, like the mature enzyme, the specificity restriction for Gly in the P1 position. Because of this narrow specificity, it is able to cleave some peptide bonds within its proregion but is incapable of hydrolysing the Asn-Leu bond (see fig. 5) that links the proregion to the mature region in this proteinase [53].

On the other hand, the presence of proregions also seems to be required for proper folding of the mature enzymes themselves. Very low levels of active proteinase expression result from the partial or total removal of upstream regions coding for prosequences [52, 66, 69]. This observation is also consistent with the fact that thermal, chemical and acid denaturation of papaya proteinases (mature forms) is irreversible in nature [70–72]. With the exception of proglycyl endopeptidase, proregions in papaya proteinases thus serve as a folding

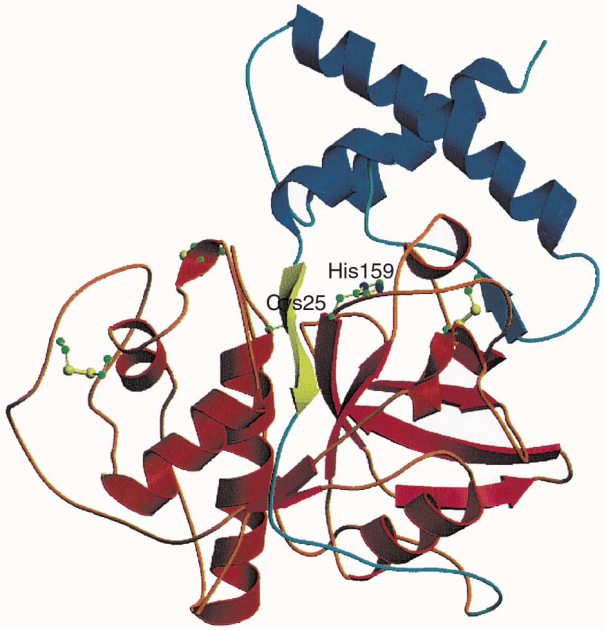


Figure 6. X-ray crystal structure of procaricain [73]. α helices are illustrated as helical ribbons and β strands as arrows. The prosequence domain is coloured in blue and the caricain enzyme in red. S-S bonds are shown in atomic detail. Positions of catalytic residues (Cys 25 and His 159) are labelled and the corresponding side chains are represented. The polypeptide segment belonging to the proregion and which binds to the active site is depicted by a yellow arrow. Obtained using Molscript [98] and Raster 3D [99] programs.

template and as an intrinsic inhibitor, preventing ectopic activation of the newly synthesised protein.

Such a dual function has gained further support through knowledge of the X-ray structure of procaricain [73] illustrated in figure 6. This structure, determined at 3.2-Å resolution, clearly shows that the proregion folds into a distinct domain. This domain consists (see also fig. 5) of three helices, a short β strand and an extended polypeptide chain that passes through the substrate-binding cleft to the N-terminus of the mature region. Several hydrogen bonds and salt bridges, but of overall importance, a hydrophobic core, stabilise this domain. The X-ray structure of procaricain also shows that the mature portion of the precursor adopts the same conformation as caricain [57], indicating that activation does not induce any substantial conformational alteration. The presence of an extended polypeptide segment belonging to the proregion and, of note, running in the opposite direction to that of substrates, prevents substrate binding. Hence, the inability of precursors to cleave even low-molecular-weight synthetic substrates is probably because the latter cannot occupy the binding cleft.

It is tempting to speculate that, in vivo, inactive papaya proteinases are stored within the laticifers [74] as their proenzyme forms. SDS-PAGE experiments support such a hypothesis, revealing, in freshly collected papaya latex only, the presence of polypeptides with molecular weights in the range (36–38 kDa) of those expected for their precursors [63, 64].

It is also tempting to hypothesise that proglycyl endopeptidase could play an important role in the in vivo processing of the proteinase precursors. The X-ray structure of procaricain strongly suggests that intramolecular cleavage in the prodomains of the four papaya proteinases is unlikely since the polypeptide

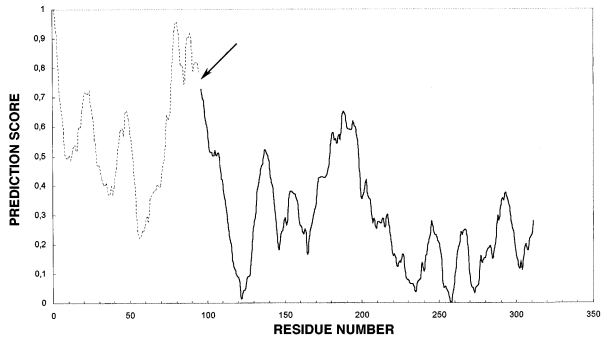


Figure 7. Prediction of the nicking sites in the procaricain structure. The program NICKPRED [105] was used to identify the peptide bonds, in the radiocrystallographic structure of procaricain [73], that are prone to proteolysis. The arrow indicates the peptide bond that delimits the proregion (dotted line) from the mature region (continuous trace).

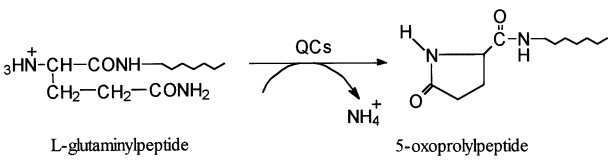


Figure 8. Schematic representation of the conversion of L-glutaminylpeptide into 5-oxoprolylpeptide as catalysed by QCs.

chains pass through the active sites in a manner reverse to that expected for substrates. Binding in the reverse direction necessary for cleavage would require the dissociation of the prodomains from the mature regions. Both the proregions of papain and papaya proteinase IV have been shown to exhibit strong affinities towards papain, chymopapain and caricain with dissociation constants in the nanomolar range [75]. Glycyl endopeptidase, on the other hand, binds much less strongly to these proregions. With dissociation constants in the micromolar range, the proform of this proteinase appears as the likely candidate for initiating the maturation process for all the papaya proteinases.

Re-examination of the procaricain structure shows (see fig. 7) that peptide bonds located within the segment p75–p90 of the proregion are good candidates for proteolysis. Interestingly, cleavage of the peptide bond between Gly p84 and Ser p85 has been observed in the course of the autocatalytic activation of the propapaya proteinase IV (Asp p1 Gly) mutant [53]. Comparison of the X-ray structures of procathepsin L and procaricain [76] provides strong evidence that the four precursors of papaya proteinases will adopt quite similar 3D folds and the invariant residue Gly p84 is thus the likely candidate as the primary site of nicking by propapaya proteinase IV. In the X-ray structure of procaricain, the O atoms of Val p83 and Gly p84 interact with the N δ 1 and the S atoms of His 159 and Cys 25, respectively, both of which are essential for endowing catalytic competence on papaya proteinases. Cleavage of the peptide bond between Gly p84 and Ser (or Phe) p85 is thus predicted to have important consequences that can lead to disencumbering the substrate-binding clefts.

In vivo, the conversion of inactive forms into fully active proteinases only takes place when latex is expelled from laticifers. In undamaged laticifers, the level of amidase activity is rather low. Following release of latex, however, the concentration in active proteases increases rapidly to reach a maximum in less than 2 min after the laticifers have been injured. At the same time, the first signs of latex coagulation also become visible. Activation of papaya proteinases is not the sole event that affects the chemical composition of coagulating latex. SDS-PAGE experiments have shown that several

other polypeptides are also proteolytically processed in an orderly fashion [63, 64].
Latex coagulation is probably the primary defence mechanism since it leads to wound sealing to protect the plant from pathogen invasion.

Papaya glutaminyl cyclotransferase

Glutaminyl (or glutamine, Q) cyclotransferases (or cyclases, QCs; EC 2.3.2.5) catalyse the conversion of N-terminal glutaminyl residues in peptides and proteins into 5-oxoprolyl (or pyroglutamyl, $<$E) residues. At physiological pHs, the α-amino function of N-terminal glutaminyl residues is protonated and cyclisation proceeds with the concomitant release of an ammonium ion (fig. 8).
In 1964, Messer and Ottesen [77] first reported the presence, in papaya latex, of an enzyme that exhibited such a catalytic activity. Early preparations of papaya QC (PQC), however, were contaminated with proteolytic activity, raising doubts about several of the physicochemical properties reported for this enzyme at that time [77, 78]. Nonetheless, these studies firmly established that interestingly, compared to several di- and hexapeptides N-terminated by glutamine, L-glutamine itself was a poor substrate for PQC. In addition, the catalytic site of the newly discovered enzyme appeared distinct from that of the papaya proteinases in its susceptibility to inhibition by several chemicals. With homogeneous preparations of PQC now available, its mechanism of action and several aspects of its structure are better known [62, 79–83].
PQC, a glycoprotein with an apparent molecular weight (SDS-PAGE) of 32–33 kDa, is a minor constituent (0.5–1%) of the protein fraction present in papaya latex. It contains a unique polypeptide chain with two free and accessible thiol functions whose conversion, e.g. as S-methylthio derivatives, does not alter its folding or its catalytic competence [81, 82]. Despite lacking disulphide bonds, PQC displays a quite remarkable resistance to chemical, thermal and acid denaturation and proteolysis. It exhibits, for example, catalytic activity over a broad pH range that extends from 3.5 to 11.0 [78, 79]. Residual native-like structure, however, still persists at more acidic pH values [83]. At neutral pH, PQC does not unfold at elevated temperature (e.g. 95 °C), or in the presence of a strong chemical denaturant (e.g. 7 M guanidinium hydrochloride) [83]. PQC is also recovered unaltered after incubation (24 h, 37 °C) in the presence of equimolar trypsin or chymotrypsin [62]. Preliminary results from SDS-PAGE experiments, also suggest that the papaya enzyme resists proteolysis by general endopeptidases such as thermolysin and subtilisin (unpublished observations).
The structural cause of this quite remarkable stability has recently been tentatively identified. Combining analysis of infrared and circular dichroism spectra with hydrogen-deuterium exchange characteristics has led to the conclusion that PQC contains an extensive β sheet structure and that its β strands are connected by short and immobile loops [62].
As a catalyst, the role of PQC primarily consists in modifying the conformation of the substrate so as to bring closer both δ carbon and α nitrogen atoms from the glutaminyl residue. By doing so, PQC facilitates the key chemical step comprising the intramolecular nucleophilic attack by the electron pair of the N atom of the α-amino group on the sp^2-hybridised C δ atom of the substrate, leading to cyclisation [79]. As expected from such a mechanism, binding of substrates to the catalytic site of PQC is the main specificity-determining factor [80].
Dahl and co-workers have recently reported the sequence of a PQC cDNA that encodes a 288-amino-acid protein including a 22-amino-acid endoplasmic reticulum signal (EMBL accession number AF061240). To the best of our knowledge, a polypeptide chain with the predicted sequence has not yet been produced by recombinant technologies nor has its ability to act on QC substrates been demonstrated. Nonetheless, the first 25 amino acids in the predicted mature region of the putative QC are fairly well matched by those constituting the N-terminal amino acid sequence of true PQC, isolated and purified from laticifers [82]. The presence of precisely two cysteinyl residues, located at positions 173 and 251 in the predicted sequence, is also quite consistent with the presence of two moles of free thiol function per mole of purified PQC [83]. Taking into account the known substrate requirement for N-glycosylation in plant cells [84], a single site may be identified, at position 101, in the predicted sequence. Assuming that the N-glycan in PQC is identical to those attached on PPI [46], glycosylated PQC can be attributed a molecular weight of 32,419 Da. This value agrees quite well with the experimental value provided by SDS-PAGE experiments [79]. The chromatographic behaviour of PQC on cation exchange supports is that of a highly basic polypeptide [82, 83]. On the other hand, a theoretical isoelectric point of 7.90 was calculated for PQC on the basis of its predicted amino acid sequence. Preferential masking of a cluster of carboxylic acid functions by the glycan moiety could provide an explanation for this apparent inconsistency.
Sequencing chromosome 4 of *Arabidopsis thaliana* led to the discovery of a QC precursor-like protein (EMBL accession number AL050400) that shares, in the expected mature region, 62% (157 out of 250) identity with the predicted sequence of PQC. Amino acid se-

quences of plant QCs show no similarity at all with those deduced from QC cDNA clones isolated from bovine and human pituitary cDNA libraries [85, 86]. Mammalian and plant QCs, therefore, appear to belong to distinct enzyme families, maybe reflecting different physiological functions.

Participation of mammalian QCs in the context of hormonal and neurotransmitter peptides biosynthesis is well established. These peptides are often synthesised as inactive precursors that are then secreted and processed. Processing requires various enzymes (e.g. peptidases, glycine α-amidating monooxygenases) including QCs [83]. It is tempting to suggest that, similar to mammalian QCs, PQC acts in concert with the papaya proteinases. Compared to mammalian tissues, however, papaya laticifers store tremendous amounts of PQC as if the plant can forecast that the physiological function of this enzyme would be required for only a short period.

Within the time interval (2 min) required for latex coagulation, one can calculate that ammonium ions are produced to a concentration of around 25 mM [79, 82], a concentration known to adversely affect the growth of micro-organisms [87]. One may further speculate that ammonium ions could also act as a chemical signal to help guide the searching behaviour of the herbivore predators and parasites [2].

The papaya lipase(s)

Lipase activity was reported in latex of *C. papaya* as early as 1935 but the enzyme responsible for this activity was only characterised more recently [88]. The papaya enzyme catalyses hydrolysis of acyl glycerols [88, 89], esterification of various fatty acids [90, 91] as well as interesterification reactions [92, 93].

As an acyl-glycerol hydrolase (EC 3.1.1.3), papaya lipase preferentially cleaves ester bonds at the *sn*-3 position [89] and shows a preference for short-chain (from C_4 to C_{10}) fatty acids [88]. Both *sn*-3 stereoselectivity and typoselectivity to short-chain fatty acids have also been observed in the course of transesterification experiments [92]. On the other hand, as a biocatalyst of esterification reactions, *C. papaya* latex strongly discriminates against fatty acids having *cis*-4, *cis*-6 and *cis*-8 unsaturations. Fatty acids with *cis*-5, *cis*-9 and *trans*-9 unsaturations are very well accepted as substrates. This is also true for fatty acids that bear hydroxy, epoxy or cyclopentenyl groups [90, 91]. The observed substrate specificities are thus similar to those reported for lipase preparations from micro-organisms and animals.

Contrasting with all the other papaya enzymes reviewed here, attempts to prepare aqueous solutions of the lipolytic enzyme have so far been unsuccessful [88, 94], indicating that the lipase activity is tightly associated with the particulate fraction of papaya latex. Obtaining preparations of papaya lipase devoid of other enzymatic activities is thereby greatly facilitated. On the other hand, however, detailed characterisation of the enzyme, at a molecular level, must await successful extraction of the polypeptide chain from its core matrix. Furthermore, its function as a defensive protein remains quite elusive.

Concluding remarks

C. papaya expresses, in its laticifers, a set of proteins and enzymes that obviously do not serve to fulfil cellular functions. Several of these proteins and enzymes have been recognised as being involved in direct defence mechanisms elaborated by plants to fend off pathogens and insects. Such is the case of glycosyl hydrolases and proteinase inhibitors [1, 2]. Taking into account the defensive role generally attributed to plant laticifers [3, 4], their presence in papaya latex is thus not surprising.

Contrasting with this situation, such a defensive role for the papaya proteinases has long been questionable. However, it now appears likely that they could indeed fulfil such a role through at least three different pathways.

First, their catalytic competence is required for latex coagulation [63, 64]. Latex coagulation seals the wound that produced its outflow, and wound sealing is probably the most efficient way to interrupt pathogens entry into the plant. The rate of coagulation is expected to depend crucially on the concentration of proteinases. Hence, expression of tremendous amounts of proteinases in papaya laticifers may now be regarded as a necessity to rapidly achieve this function.

We might further speculate that remodelling the polypeptide chains expressed in papaya laticifers through proteolysis can also serve quite different defence strategies more specific for *Carica* species.

In concert with glutaminyl cyclase, proteolysis can generate ammonium ion production that, theoretically, could affect the growth of micro-organisms. The list of peptides known to have a direct toxic action on micro-organisms is growing rapidly. Such peptides can be generated from several unrelated proteins and enzymes, including lysozymes [32] and cystatins [47]. Production of toxic doses of ammonium ions and of microbicidal peptides by *Carica* species would constitute an efficient means to respond to adaptation by insects [2] and pathogens [1, 95]. Experimental evidence that toxic ammonium ions and microbicidal peptides are produced when latex is expelled from the papaya plant is still lacking. However, papaya latex is well equipped for

these functions. Both then, urgently deserve further attention.

The above-suggested synergy between proteinases and glutaminyl cyclase to produce ammonium ions may not be unique. Some other observations also suggest that collaboration between the proteinases themselves is likely. These enzymes are stored in their inactive precursor forms that become activated when latex is expelled. Our actual knowledge of the mechanism of the activation process is solely based upon results from in vitro experiments carried out on individual proteinases. This may be the reason why activation of the precursors after plant damage proceeds much more rapidly than expected on the basis of these in vitro experiments.

However, the in vivo situation can rarely be reproduced exactly. Considering just the proteinase family, no less than nine molecular species have been identified thus far (see table 1). Some of their precursors have N-glycosylation sites, and others, C-terminal extensions. C-terminal extensions are known as plant vacuolar sorting signals [96] and, of note, the correct processing of at least one other member of the papain superfamily has been shown in the past to require the presence of such an extension [97]. Hence, examination of the predicted amino acid sequences of the precursors indicates that several sites of sequestration may be provided for these proteinases. In vivo trafficking in the papaya laticifers is another unexplored facet of the physiology of this plant.

Thus, although several important actors implicated in the direct defence mechanisms of *C. papaya* have now been identified, much research effort must still be devoted to acquire a better understanding of the complex ecophysiological responses provided by this plant.

1 Somssich I. E. and Halbrock K. (1998) Pathogen defence in plants: a paradigm of biological complexity. Trends Plant Sci. **3:** 86–90

2 Baldwin I. T. and Preston C. A. (1999) The eco-physiological complexity of plant responses to insect herbivores. Planta **208:** 137–145

3 Dussourd D. E. and Eisner T. (1987) Vein-cutting behaviour: insect counterplay to the latex defense of plants. Science **237:** 898–901

4 Malcolm S. B. and Zalucki M. P. (1996) Milkweed latex and cardenolide induction may resolve the lethal plant defense paradox. Entomol. Exp. Appl. **80:** 193–196

5 Jacob J. L., Prevôt J. C. and Kekwick R. G. O. (1989) General metabolism of *Hevea brasiliensis* latex. In: Physiology of Rubber Tree Latex, pp. 145–178, d'Auzac J., Jacob J. L. and Chrestin H. (eds), CRC, Boca Raton

6 Dubois T., Jacquet A., Schnek A. G. and Looze Y. (1988) The thiol proteinases from the latex of *Carica papaya* L. I. Fractionation, purification and preliminary characterization. Biol. Chem. Hoppe-Seyler **369:** 733–740

7 Walraevens V., Jaziri M., Van Beeumen J., Schnek A. G., Kleinschmidt T. and Looze Y. (1993) Isolation and preliminary characterization of the cysteine proteinases from the latex of *Carica candamarcensis*. Hook. Biol. Chem. Hoppe-Seyler **374:** 501–506

8 Roth I. and Clausnitzer I. (1972) Desarrollo y anatomia del fruto y de la semilla de *Carica papaya* L. (Lechosa). Acta Bot. Venez. **7:** 187–206

9 Mitchel R. E. J., Chaiken I. M. and Smith E. L. (1970) The complete amino acid sequence of papain: additions and corrections. J. Biol. Chem. **245:** 3485–3492

10 Drenth J., Jansonius J. N., Koekoek R., Swen H. M. and Wolthers B. G. (1968) Structure of papain. Nature **218:** 929–932

11 Brocklehurst K. and Salih E. (1983) A re-evaluation of the nomenclature of the cysteine proteinases of *Carica papaya* and a rational basis for their identification. Biochem. J. **213:** 559–560

12 Polgar L. (1984) Problems of classification of papaya latex proteinases. Biochem. J. **221:** 555–556

13 Barrett A. J. and Buttle D. J. (1985) Names and number of papaya proteinases. Biochem. J. **228:** 527

14 Buttle D. J., Kembhavi A. A., Sharp S. L., Shute R. E., Rich D. H. and Barrett A. J. (1989) Affinity purification of the novel cysteine proteinase papaya proteinase IV and papain from papaya latex. Biochem. J. **261:** 469–476

15 Stinzi A., Heitz T., Prasad V., Wiedemann-Merdinoglu S., Kauffmann S., Geoffroy P. et al. (1993) Plant pathogenesis-related proteins and their role in defense against pathogens. Biochimie **75:** 687–706

16 Wilson G. (1974) Some properties of an endo-1,3-β-glucanase. Biochem. Soc. Trans. **2:** 1115–1116

17 Smith E. L., Kimmel J. R., Brown D. M. and Thompson E. O. P. (1955) Isolation and properties of a crystalline mercury derivative of a lysozyme from papaya latex. J. Biol. Chem. **215:** 67–89

18 Howard J. B. and Glazer A. N. (1967) Studies of the physicochemical and enzymatic properties of papaya lysozyme. J. Biol. Chem. **242:** 5715–5723

19 Howard J. B. and Glazer A. N. (1969) Papaya lysozyme: terminal sequences and enzymatic properties. J. Biol. Chem. **244:** 1399–1409

20 Dahlquist F. W., Borders C. L. Jr, Jacobson G. and Raftery M. A. (1969) The stereospecificity of human, hen and papaya lysozymes. Biochemistry **8:** 694–700

21 Barel A. O., Dolmans M. and Léonis J. (1971) Spectroscopic studies on the reduction and reformation of the disulfide bonds of papaya lysozyme. Eur. J. Biochem. **19:** 488–495

22 Heneine I. F. and Kimmel J. R. (1972) The sulfhydryl groups of papaya lysozyme and their relation to biological activity. J. Biol. Chem. **247:** 6589–6596

23 Henrissat B. and Bairoch A. (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **293:** 781–788

24 Beintema J. J. (1994) Structural features of plant chitinases and chitin-binding proteins. FEBS Lett. **350:** 159–163

25 Azarkan M., Amrani A., Nijs M., Vandermeers A., Zerhouni S., Smolders N. et al. (1997) *Carica papaya* latex is a rich source of a class II chitinase. Phytochemistry **46:** 1319–1325

26 Hart P. J., Pfluger H. D., Monzingo A. F., Hollis T. and Robertus J. D. (1995) The refined crystal structure of an endochitinase from *Hordeum vulgare* L. seeds at 1.8 Å resolution. J. Mol. Biol. **248:** 402–413

27 Terwisscha Van Scheltinga A. C., Hennig M. and Dijkstra B. W. (1996) The 1.8 Å resolution structure of hevamine, a plant chitinase/lysozyme, and analysis of the conserved sequence and structure motifs of glycosyl hydrolase family 18. J. Mol. Biol. **262:** 243–257

28 Subroto T., Sufiati S. and Beintema J. J. (2000) Papaya (*Carica papaya*) lysozyme is a member of the family 19 (basic, class II) chitinases. J. Mol. Evol. **49:** 819–821

29 Bokma E., Koningsveld G. A. van, Jeronimus-Stratingh M. and Beintema J. J. (1997) Hevamine, a chitinase from the rubber tree *Hevea brasiliensis*, cleaves peptidoglycan between the C-1 of N-acetylglucosamine and C-4 of N-acetylmuramic acid and therefore is not a lysozyme. FEBS Lett. **411:** 161–163

30 Iseli B., Armand S., Boller T., Neuhaus J. M. and Henrissat B. (1996) Plant chitinases use two different hydrolytic mechanisms. FEBS Lett. **382:** 186–188

31 Terwisscha Van Scheltinga A. C., Armand S., Kalk K. H., Isogai A., Henrissat B. and Dijkstra B. W. (1995) Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and X-ray structure of a complex with allasamidin: evidence for substrate assisted catalysis. Biochemistry **34:** 15619–15623

32 Düring K., Porsch P., Mahn A., Brinkmann O. and Gieffers W. (1999) The non-enzymatic microbicidal activity of lysozymes. FEBS Lett. **449:** 93–100

33 Song I., Taylor M., Baker K. and Bateman R. C. Jr (1995) Inhibition of cysteine proteinases by *Carica papaya* cystatin produced in *Escherichia coli*. Gene **162:** 221–224

34 Rogelj B., Popovic T., Ritonja A., Strukelj B. and Brzin J. (1998) Chelidocystatin, a novel phytocystatin from *Chelidonium majus*. Phytochemistry **49:** 1645–1649

35 Kondo H., Abe K., Emori Y. and Arai S. (1991) Gene organization of orizacystatin-II, a new cystatin superfamily member of plant origin, is closely related to that of orizacystatin-I but different from those of animal cystatins. FEBS Lett. **278:** 87–90

36 Ryan S. N., Laing W. A. and McMannus M. T. (1998) A cysteine proteinase inhibitor purified from apple fruit. Phytochemistry **49:** 957–963

37 Turk V. and Bode W. (1991) The cystatins: protein inhibitors of cysteine proteinases. FEBS Lett. **285:** 213–219

38 Bjork I. and Ylinenjärvi K. (1989) Interaction of chicken cystatins with inactivated papains. Biochem. J. **260:** 61–68

39 Rowan A. D., Brzin J., Buttle D. J. and Barrett A. J. (1990) Inhibition of cysteine proteinases by a protein inhibitor from potato. Biochem. J. **269:** 328–330

40 Pernas M., Sanchez-Monge R., Gomez L. and Salcedo G. (1998) A chestnut seed cystatin differentially effective against cysteine proteinases from closely related pests. Plant Mol. Biol. **38:** 1235–1242

41 Bode W., Engh R., Musil D., Theile U., Huber R., Karshikov A. et al. (1988) The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its mode of interaction with cysteine proteinases. EMBO J. **7:** 2593–2599

42 Stubbs T., Laber B., Bode W., Huber R., Jerala R., Lenarcic B. et al. (1990) The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. EMBO J. **9:** 1939–1947

43 Arai S., Watanabe H., Kondo H., Emori Y. and Abe K. (1991) Papain inhibitory activity of oryzacystatin, a rice seed cysteine proteinase inhibitor, depends on the central Gln-Val-Val-Ala-Gly region conserved among cystatin superfamily members. J. Biochem. **109:** 294–298

44 Pol E. and Bjork I. (1999) Importance of the second binding loop and the C-terminal end of cystatin B (stefin B) for inhibition of cysteine proteinases. Biochemistry **38:** 10519–10526

45 Odani S., Yokokawa Y., Takeda H., Abe S. and Odani S. (1996) The primary structure and characterization of carbohydrate chains of the extracellular glycoprotein proteinase inhibitor from latex of *Carica papaya*. Eur. J. Biochem. **241:** 77–82

46 Shimazaki A., Makino Y., Omichi K., Odani S. and Hase S. (1999) A new sugar chain of the proteinase inhibitor from latex of *Carica papaya*. J. Biochem. **125:** 560–565

47 Blankenvoorde M. F. J., 't Hof W. van, Walgreen-Weterings E., Steenbergen T. J. M. van, Brand H. S., Veerman E. C. I. et al. (1998) Cystatin and cystatin-derived peptides have antibacterial activities against the pathogen *Porphyromonas gingivalis*. Biol. Chem. **379:** 1371–1375

48 Dubois T., Kleinschmidt T., Schnek A. G., Looze Y. and Braunitzer G. (1988) The thiol proteinases from the latex of *Carica papaya* L. II. The primary structure of proteinase omega. Biol. Chem. Hoppe-Seyler **369:** 741–754

49 Jacquet A., Kleinschmidt T., Schnek A. G., Looze Y. and Braunitzer G. (1989) The thiol proteinases from the latex of *Carica papaya* L. III. The primary structure of chymopapain. Biol. Chem. Hoppe-Seyler **370:** 425–434

50 Ritonja A., Buttle D. J., Rawlings N. D., Turk V. and Barrett A. J. (1989) Papaya proteinase IV amino acid sequence. FEBS Lett. **258:** 109–112

51 Cohen L. W., Coghlan V. M. and Dihel L. C. (1986) Cloning and sequencing of papain-encoding cDNA. Gene **48:** 219–227

52 Revell D. F., Cummings N. J., Baker K. C., Collins M. E., Taylor M. A. J., Sumner I. G. et al. (1993) Nucleotide sequence and expression in *Escherichia coli* of cDNAs encoding papaya proteinase omega from *Carica papaya*. Gene **127:** 221–225

53 Baker K. C., Taylor M. A. J., Cummings N. J., Tuñon M. A., Worboys K. A. and Connerton I. F. (1996) Autocatalytic processing of propapaya proteinase IV is prevented by crowding of the active-site cleft. Protein Eng. **9:** 525–529

54 Taylor M. A. J., Al-seikh M., Revell D. F., Sumner I. G. and Connerton I. F. (1999) cDNA cloning and expression of *Carica papaya* prochymopapain isoforms in *Escherichia coli*. Plant Sci. **145:** 41–47

55 Watson D. C., Yaguchi M. and Lynn K. R. (1990) The amino acid sequence of chymopapain from *Carica papaya*. Biochem. J. **266:** 75–81

56 Kamphuis I. G., Kalk K. H., Swarte M. B. A. and Drenth J. (1984) Structure of papain at 1.65 Å resolution. J. Mol. Biol. **179:** 233–256

57 Pickersgill R. W., Rizkallah P., Harris G. W. and Goodenough P. W. (1991) Determination of the structure of papaya proteinase omega. Acta Cryst. B **47:** 766–771

58 O'Hara B. P., Hemmings A. M., Buttle D. J. and Pearl L. H. (1995) Crystal structure of glycylendopeptidase: a cysteine endopeptidase of unusual substrate specificity. Biochemistry **34:** 13190–13195

59 Maes D., Bouckaert J., Poortmans F., Wyns L. and Looze Y. (1996) Structure of chymopapain at 1.7 Å resolution. Biochemistry **35:** 16292–16298

60 Buttle D. J. (1994) Glycyl endopeptidase. Methods Enzymol. **244:** 539–555

61 Jacquet A., Kleinschmidt T., Dubois T., Schnek A. G., Looze Y. and Braunitzer G. (1989) The thiol proteinases from the latex of *Carica papaya* L. IV. Proteolytic specificities of chymopapain and papaya proteinase omega determined by digestion of alpha-globin chains. Biol. Chem. Hoppe-Seyler **370:** 819–829

62 Oberg K. A., Ruysschaert J. M., Azarkan M., Smolders N., Zerhouni S., Wintjens R. et al. (1998) Papaya glutamine cyclase, a plant enzyme highly resistant to proteolysis, adopts an all-β conformation. Eur. J. Biochem. **258:** 214–222

63 Silva L. G., Garcia O., Lopes M. T. P. and Salas C. E. (1997) Changes in protein profile during coagulation of latex from *Carica papaya*. Braz. J. Med. Biol. Res. **30:** 615–619

64 Moutim V., Silva L. G., Lopes M. T. P., Wilson Fernandes G. and Salas C. E. (1999) Spontaneous processing of peptides during coagulation of latex from *Carica papaya*. Plant Sci. **142:** 115–121

65 Vernet T., Tessier D. C., Richardson C., Laliberté F., Khouri H. E., Bell A. W. et al. (1990) Secretion of functional papain precursor from insect cells: requirement for N-glycosylation of the pro-region. J. Biol. Chem. **265:** 16661–16666

66 Taylor M. A., Pratt K. A., Revell D. F., Baker K. C., Sumner I. G. and Goodenough P. W. (1992) Active papain renatured and processed from insoluble recombinant propapain expressed in *Escherichia coli*. Protein Eng. **5:** 455–459

67 Vernet T., Chatellier J., Tessier D. C. and Thomas D. Y. (1993) Expression of functional papain precursor in *Saccharomyces cerevisiae*: rapid screening of mutants. Protein Eng. **6:** 213–219

68 Vernet T., Berti P. J., Montigny C. de, Musil R., Tessier D.C., Menard R. et al. (1995) Processing of the papain

precursor: the ionization state of a conserved amino acid motif within the pro-region participates in the regulation of intramolecular processing. J. Biol. Chem. **270:** 10838–10846
69 Cohen L. W., Fluharty C. and Dihel L. C. (1990) Synthesis of papain in *Escherichia coli*. Gene **88:** 263–267
70 Tiktopulo E. I. and Privalov P. L. (1978) Papain denaturation is not a two-state transition. FEBS Lett. **91:** 57–58
71 Solis-Mendiola S., Guttierres-Gonzales L. H., Arroyo-Reyna A., Padilla-Zuniga J., Rojo-Dominguez A. and Hernandez-Arana A. (1998) pH dependence of the activation parameters for chymopapain unfolding: influence of ion pairs on the kinetic stability of proteins. Biochim. Biophys. Acta **1388:** 363–372
72 Lopez-Arenas L., Solis-Mendiola S. and Hernandez-Arana A. (1999) Estimating the degree of expansion in the transition state for protein unfolding: pH dependence of the rate constant for caricain denaturation. Biochemistry **38:** 15936–15943
73 Groves M. R., Taylor M. A. J., Scott M., Cummings N. J., Pickersgill R. W. and Jenkins J. A. (1996) The prosequence of procaricain forms an α-helical domain that prevents access to the substrate-binding cleft. Structure **4:** 1193–1203
74 Smith H., Mc Kee R. A. and Prackelt U. M. (1987) The molecular biology of plant thiol proteinases. In: Chemical Aspects of Food Enzymes, Special Publication 63, pp. 196–207, Royal Society of Chemistry, London
75 Taylor M. A. J., Baker K. C., Briggs G. S., Connerton I. F., Cummings N. J., Pratt K. A. et al. (1995) Recombinant pro-regions from papain and papaya proteinase IV are selective high affinity inhibitors of the mature papaya enzymes. Protein Eng. **8:** 59–62
76 Groves M. R., Coulombe R., Jenkins J. and Cygler M. (1998) Structural basis for specificity of papain-like cysteine protease proregions toward their cognate enzymes. Proteins **32:** 504–514
77 Messer M. and Ottesen M. (1964) Isolation and properties of glutamine cyclotransferase of dried papaya latex. Biochim. Biophys. Acta **92:** 409–411
78 Messer M. and Ottesen M. (1965) Isolation and properties of glutamine cyclotransferase of dried papaya latex. C. R. Trav. Lab. Carlsberg **35:** 1–24
79 Gololobov M. Y., Song I., Wang W. and Bateman R. C. Jr (1994) Steady-state kinetics of glutamine cyclotransferase. Arch. Biochem. Biophys. **309:** 300–307
80 Gololobov M. Y., Wang W. and Bateman R. C. Jr (1996) Substrate and inhibitor specificity of glutamine cyclotransferase (QC). Biol. Chem. Hoppe-Seyler **377:** 395–398
81 Amrani A., Wintjens R., Zerhouni S., Azarkan M., Nijs M., Vincentelli J. et al. (1997) The conversion of N-terminal glutaminyl residues of plant proteins into 5-oxoproline: a post-translational modification in quest of a physiological function. Recent Res. Dev. Phytochem. **1:** 25–37
82 Zerhouni S., Amrani A., Nijs M., Vandermeers A. and Looze Y. (1997) Purification and characterization of the plant glutamyl-peptide cyclotransferase isolated from papaya latex. Int. J. Bio-Chromatogr. **3:** 189–206
83 Zerhouni S., Amrani A., Nijs M., Smolders N., Azarkan M., Vincentelli J. et al. (1998) Purification and characterization of papaya glutamine cyclotransferase, a plant enzyme highly resistant to chemical, acid and thermal denaturation. Biochim. Biophys. Acta **1387:** 275–290
84 Ueda H. and Ogawa H. (1999) Glycobiology of the plant glycoprotein epitope: structure, immunogenicity and allergenicity of plant glycotopes. Trends Glycosci. Glycotechnol. **62:** 413–428
85 Pohl T., Zimmer M., Mugele K. and Spiess J. (1991) Primary structure and functional expression of glutaminyl cyclase. Proc. Natl. Acad. Sci. USA **88:** 10059–10063
86 Song I., Chuang C. Z. and Bateman R. C. Jr (1994) Molecular cloning, sequence analysis and expression of human pituitary glutaminyl cyclase. J. Mol. Endocrinol. **13:** 77–86
87 Banuelos O., Casqueiro J., Gutierrez S., Riano J. and Martin J. F. (2000) The specific transport system for lysine is fully inhibited by ammonium in *Penicillium chrysogenum*: an ammonium-insensitive system allows uptake in carbon-starved cells. Antonie Van Leeuwenhoek **77:** 91–100
88 Giordani R., Moulin A. and Verger R. (1991) Tributyroylglycerol hydrolase activity in *Carica papaya* and other latices. Phytochemistry **30:** 1069–1072
89 Villeneuve P., Pina M., Montet D. and Graille J. (1995) *Carica papaya* latex lipase: sn-3 stereoselectivity or short-chain selectivity? Model chiral triglycerides are removing the ambiguity. J. Am. Oil Chem. Soc. **72:** 753–755
90 Mukherjee K. D. and Kiewitt I. (1996) Specificity of *Carica papaya* latex as biocatalyst in the esterification of fatty acids with 1-butanol. J. Agric. Food Chem. **44:** 1948–1952
91 Borgdorf R. and Warwel S. (1999) Substrate selectivity of various lipases in the esterification of *cis*- and *trans*-9-octadecenoic acid. Appl. Microbiol. Biotechnol. **51:** 480–485
92 Villeneuve P., Pina M., Skarbek A., Graille J. and Foglia T. A. (1997) Specificity of *Carica papaya* latex in lipase-catalyzed interesterification reactions. Biotechnol. Tech. **11:** 91–94
93 Mukherjee K. D. and Kiewitt I. (1998) Structured triacylglycerols resembling human milk fat by transesterification catalyzed by papaya (*Carica papaya*) latex. Biotechnol. Lett. **20:** 613–616
94 Mukherjee K. D. and Kiewitt I. (1998) Substrate specificity of lipases in protease preparations. J. Agric. Food Chem. **46:** 2427–2432
95 Bishop J. G., Dean A. M. and Mitchell-Olds T. (2000) Rapid evolution in plant chitinases: molecular targets selection in plant-pathogen coevolution. Proc. Natl. Acad. Sci. USA **97:** 5322–5327
96 Chrispeels M. J. and Raikhel N. V. (1992) Short peptide domains target proteins to plant vacuoles. Cell **68:** 613–616
97 Paul W., Amiss J., Try R., Praekelt U., Scott R. and Smith H. (1995) Correct processing of the kiwifruit protease actinidin in transgenic tobacco requires the presence of the C-terminal propeptide. Plant Physiol. **108:** 261–268
98 Kraulis P. J. (1991) Mol Script: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. **24:** 946–950
99 Merrit E. A. and Murphy M. E. P. (1994) Raster 3D version 2.0: a program for photorealistic molecular graphics. Acta Crystallogr. D **50:** 869–873
100 Sali A. and Blundell T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. **234:** 779–815
101 Song H. K. and Suh S. W. (1998) Kunitz-type soybean trypsin inhibitor revisited: refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from *Erythrina caffra* and tissue-type plasminogen activator. J. Mol. Biol. **275:** 347–363
102 Onesti S., Brick P. and Blow D. M. (1991) Crystal structure of a Kunitz-type trypsin inhibitor from *Erythrina caffra* seeds. J. Mol. Biol. **217:** 153–176
103 Dayan S. M., Van Donkelaar A. and Kort A. A. (1987) Crystallographic data of the major albumin from *Psophocarpus tetragonolobus* (L.) DC. J. Biol. Chem. **262:** 10287–10289
104 Dattagupta J. K., Chakrabarti C., Podder A., Dutta S. K. and Singh H. (1990) Crystallization and preliminary X-ray studies of psophocarpin B1, a chymotrypsin inhibitor from winged bean seeds. J. Mol. Biol. **216:** 229–231
105 Hubbard S. J., Beynon R. J. and Thornton J. M. (1998) Assessment of conformational parameters as predictors of limited proteolytic sites in native protein structures. Protein Eng. **11:** 349–359