

Structural, Biological, and Evolutionary Relationships of Plant Food Allergens Sensitizing via the Gastrointestinal Tract

E. N. CLARE MILLS and JOHN A. JENKINS

Institute of Food Research, Norwich Research Park, Colney, Norwich, United Kingdom

MARCOS J. C. ALCOCER

School of Biological Sciences, University of Nottingham, Nottingham, United Kingdom

PETER R. SHEWRY

IACR-Long Ashton Research Station, Rothamsted Research, Rothamsted, Harpenden, United Kingdom

The recently completed genome sequence of the model plant species Arabidopsis has been estimated to encode over 25,000 proteins, which, on the basis of their function, can be classified into structural and metabolic (the vast majority of plant proteins), protective proteins, which defend a plant against invasion by pathogens or feeding by pests, and storage proteins, which proved a nutrient store to support germination in seeds. It is now clear that almost all plant food allergens are either protective or storage proteins. It is also becoming evident that those proteins that trigger the development of an allergic response through the gastrointestinal tract belong primarily to two large protein superfamilies: (1) The cereal prolamin superfamily, comprising three major groups of plant food allergens, the 2S albumins, lipid transfer proteins, and cereal α -amylase/trypsin inhibitors, which have related structures, and are stable to thermal processing and proteolysis. They include major allergens from Brazil nut, peanuts, fruits, such as peaches, and cereals, such as rice and wheat; (2) The cupin superfamily, comprising the major globulin storage proteins from a number of plant species. The globulins have been found to be allergens in plant foods, such as peanuts, soya bean, and walnut; (3) The cysteine protease C1 family, comprising the papain-like proteases from microbes, plants, and animals. This family contains two notable allergens that sensitize via the GI tract, namely actinidin from kiwi fruit and the soybean allergen, Gly m Bd 30k/P34. This study describes the properties, structures, and evolutionary relationships of these protein families, the allergens that belong to them, and discusses them in relation to the role protein structure may play in determining protein allergenicity.

Keywords cupins, cysteine, prolamins, protease, protein families, structure

INTRODUCTION

The term allergy was first coined by von Pirquet, in 1906 and referred to the 'changed reactivity' of an individual following exposure to an 'agent' for a second or subsequent occasion. Such allergic reactions (also known as hypersensitivity reactions) are now known to occur following a variety of environmental challenges, including foods. They are a consequence of inappropriate immune responses that may result in inflammatory

reactions and tissue damage and were classified many years ago by Coombs and Gell on a symptomatic basis into Types I, II, III, and IV. Of these, only Type I hypersensitivity reactions involve generation of an IgE antibody response toward an environmental agent, the IgE then binding to the surface of histamine containing mast cells. Following exposure to the same environmental agent on a subsequent occasion, the agent becomes bound to several IgE molecules, cross-linking them. This event triggers the release of mediators, such as histamine from the mast cells, which go on to trigger an acute inflammatory reaction manifested as respiratory (asthma, rhinitis), skin (eczema, urticaria), or gastrointestinal symptoms (vomiting, diarrhoea), either alone or in combination. A rare, but very severe, reaction is anaphylactic

Address correspondence to E. N. Clare Mills, Institute of Food Research, Norwich Research Park, Colney, Norwich, NR4 7UA, UK. E-mail: clare.mills@BBSRC.AC.UK

shock characterized by respiratory symptoms, fainting, itching, urticaria, swelling of the throat or other mucous membranes, and a dramatic loss of blood pressure.

The macromolecules able to sensitize an individual and subsequently elicit an allergic reaction are called allergens, and they are almost always proteinaceous in nature. In Type I hypersensitivity reactions, this usually means that an allergen is the protein toward which an IgE response is generated. However, not all IgE causes an allergic response, as has been found for peanut allergic patients who have serum IgE that cross-reacts with soy proteins, but have no demonstrable allergic reactions on consuming soy (Eigenmann et al., 1996). Allergens can also be classified as being either major, i.e., one recognized by IgE from 50% of a group of at least ten allergic patients, or minor (Larsen and Lowenstein, 1996). However, such a designation does not carry any connotation of allergenic potency and is disputed by some (Alberse, 2000).

The normal function of IgE is to provide some protection against parasitic infestations, but the mechanism involved in triggering its production in allergic disease has not been defined. It is clear that factors such as the level of exposure to allergens and genetic predisposition play an important role. Sensitization may occur via the skin, inhalation, or the ingestion of substances. Other factors, such as respiratory tract viral infections and environmental pollutants, may either reduce or enhance allergen-specific IgE production, possibly by increasing mucosal permeability in organs, such as the lungs (Hopkin 2002; Salvi, 2001). Food allergy sensitization usually occurs via the gastrointestinal tract and frequently results in systemic reactions. There are other food allergy syndromes where sensitivity to a food is associated with prior sensitization to an inhalant allergen, such as pollen or latex. In such instances, IgE developed to pollen or latex allergen cross-reacts with homologous proteins found in foods, such as fruits and vegetables (Breiteneder and Ebner, 2000). In general, symptoms from these cross-reacting allergies tend to be less severe and are confined to the mouth, with itching and swelling around the lips, leading to the term 'Oral Allergy Syndrome' (OAS), although in some instances reactions may be more severe, even life-threatening (Anderson, 1996).

PLANT FOOD ALLERGIES

Along with other types of allergic disease, the prevalence of food allergy appears to be increasing. Thus, 1–2% of adults are thought to suffer from some form of food allergy, the prevalence being higher in children (up to 7%) (Anderson, 1996). While for a large proportion of sufferers the allergic reactions are debilitating, for some they can be life threatening. In particular, the incidence of more severe reactions, such as anaphylaxis, toward foods, such as peanuts seems to be rising, particularly in children. Indeed, peanuts are one of most allergenic foods, frequently causing very severe reactions. For example, it has been estimated that 28% of food allergies in France arise from peanuts (Moneret-Vautrin et al., 1998), with almost all the allergic sub-

jects being under 7 years of age. Another cause for concern is that unlike allergies to cow's milk, where almost all children lose their sensitivity by the age of 7, allergy to some foods, such as peanuts, appears to be life long.

A wide range of plant foods have been cited as causing allergic reactions, and of the seven main foods that are responsible for causing the majority of food allergies, four are of plant origin, namely wheat, peanuts, soy, and tree nuts (Bush and Hefle, 1996). Currently, we do not understand what makes certain foods more allergenic than others, but the level of exposure to sensitizing agents undoubtedly plays a part. Thus, the distribution of allergies differs between countries depending on food consumption, with buckwheat allergy being much more common in the Far East where it is widely consumed, than is the case in Western countries (Park et al., 2000). Patterns of cross-reacting allergies also differ geographically, as related to the presence of both the environmental sensitizing agent and consumption of fruits and vegetables. In areas where birch pollen allergy is common, there is a higher incidence of associated apple allergy (Pauli et al., 1996), while in central Europe, allergies to birch and mugwort pollen result in reactions to vegetables, such as celery (Bauer et al., 1996).

In addition to the levels of exposure to allergens, it is possible that certain structural and biological properties of proteins predispose them to becoming allergens by increasing their uptake and interactions with the immune system. In order to elicit an IgE response, an allergen must be able to enter the body (i.e., cross a mucosal barrier) and then survive *in vivo*, while retaining sufficient structural integrity for a sufficient length of time to interact with the immune system. Thus, the structural features of a protein that predispose it to become an allergen will be largely determined by the route of sensitization. Hence, food allergens sensitizing via the gastrointestinal tract must be stable to thermal processing and resistant to the digestive process, while inhalant allergens must be readily soluble in physiological buffers. Many possess proteolytic activity, enabling the allergen to breach the respiratory mucosa (Stewart and Thompson, 1996). A fuller understanding of the role that protein structure plays in allergenicity is a necessity, as regulatory procedures require an assessment of the allergenic properties of a novel food (including genetically modified organisms [GMOs] destined for food use), before they can be approved for human consumption (Kimber et al., 1999). Indeed, a WHO-FAO consultation on the allergenicity of foods derived from biotechnology suggested that the allergenic risk assessment process should take into account whether a candidate transgene encodes a protein belonging to a known plant allergen protein family (FAO-WHO, 2001). It is also important if *in vitro* diagnostics of allergic disease are to be improved and will pave the way for development of therapeutic strategies, to use modified allergens to desensitize individuals.

This review discusses the comparative structures and properties of plant food allergens, primarily those found in seeds, which are thought to sensitize individuals via the gastrointestinal tract, including common structure and properties that may relate to their allergenic potential. The properties of pollen and

latex-fruit/vegetable allergens have been described elsewhere (Breiteneder and Ebner, 2000; Hoffman-Sommergruber, 2000).

PLANT AND SEED PROTEINS

The study of plant proteins has a long history, extending to 1745, when the isolation of wheat gluten was first described (Beccari, 1745). However, the first systematic study of plant proteins was carried out by T.B. Osborne, while working at the Connecticut Agricultural Experiment Station (USA) between 1886 and 1928. He published studies of proteins from seeds of 32 plant species leading to his monograph "The Vegetable Proteins," published in 1924 (Osborne, 1924). Previous workers had used a range of solvents to extract and separate plant proteins, and Osborne combined these into a sequential procedure, which is now generally called "Osborne fractionation".

This procedure is designed to be applied to milled seed material which is extracted sequentially with water to extract albumins, dilute saline (0.5–1.0 M NaCl) to extract globulins, alcohol-water mixtures (traditionally 60–70% (v/v) ethanol) to extract prolamins and dilute acid, or alkali to extract glutelins, a small proportion of the total proteins remaining in the residue. This procedure has been modified in a number of ways since first introduced, particularly by the replacement of ethanol with other alcohols (e.g., 50% (v/v) propan-1-ol) and the inclusion of a reducing agent to extract prolamin subunits present in polymers associated by interchain disulphide bonds, and by the extraction of the glutelins with detergents (e.g., SDS) and/or chaotropic agents (e.g., urea), instead of acids or alkalis, which can cause partial degradation. Although these modifications have led to the production of purer and better defined fractions, it still suffers from the drawback that the same Osborne fraction may contain different types of proteins in the same seed or in different seeds or other plant tissues.

Thus, seed storage proteins may be present in the albumin and globulin fractions, where they are mixed with smaller amounts of non-storage proteins, while the prolamins form a well-defined storage protein fraction in cereal grain, but contain only minor non-storage components in other species. The glutelins usually comprise insoluble structural and metabolic proteins, but also contain residual globulin and prolamin storage proteins that have not been extracted by the previous solvents. Furthermore, the Osborne fractions have little meaning when applied to vegetative tissues in which storage proteins are not present. Although many seed biologists still use the Osborne classification, there is now a consensus of opinion that it is more valid to initially divide plant proteins into three broad groups: structural and metabolic, protective, and storage.

CLASSIFICATION OF PLANT PROTEINS

Structural and Metabolic Proteins

These include a wide range of proteins, which play roles in the structure and function of plant cells. Many may be regarded

as "housekeeping" proteins, since they are essential for the structure and function of all cells, while others are synthesized only in specific cell types or tissues at a specific developmental stage or in response to specific signals. They include enzymes involved in biosynthesis and catabolism, structural proteins present in membranes and cell walls, transporters, and components of signalling cascades. For example, analysis of the proteins encoded by the *Arabidopsis* genome (a model plant species related to crucifers such as cabbage and oilseed rape) predicts that about 22.5% of the genes encode enzymes of primary and secondary metabolism and 11% encode proteins involved in defence, while the remainder encode proteins involved in growth and division, energy generation, signal transduction, transcription, protein synthesis, and cellular trafficking and transport (The Arabidopsis Genome Initiative, 2000).

Most structural and metabolic proteins are present in very small amounts, sometimes too small to be identified by electrophoresis of total protein fractions (for example, components of signalling pathways). However, this is not the case for some other components, such as photosynthetic enzymes and associated proteins in green tissues. The most abundant of these proteins, ribulose-1, 5-bisphosphate carboxylase/oxygenase (usually abbreviated to rubisco), accounts for about 30–40% of total leaf protein in most species.

Despite the widespread consumption of vegetative tissues by humans, there is little evidence that either the major photosynthetic enzymes or the less abundant structural and metabolic proteins provoke allergic reactions.

Protective PR Proteins

Plant tissues contain a range of proteins and peptides that have been demonstrated to inhibit the growth of pathogenic microorganisms (fungi, bacteria, viruses) or the feeding or infection by invertebrate pests (e.g., nematodes, molluscs, insects), either *in vivo* (e.g., in transgenic plants) or *in vitro* (e.g., in feeding tests). These proteins are widely considered to play a natural role in plant resistance, although this is difficult to establish conclusively. They are also considered to have potential for exploitation to confer resistance in transgenic crops and a number of putative protective proteins have been demonstrated to have such effects when tested in model crop plants or under controlled growth conditions (Shewry and Lucas, 1997). However, despite the fact that this effect was first described over a decade ago (Hilder et al., 1987), commercial crops expressing plant-derived protective proteins have yet to be produced. This contrasts with the success achieved using the endotoxins derived from the bacterium *Bacillus thuringiensis* (Bt toxins) in crops such as maize (corn) and cotton.

The most widely studied plant defense response is the pathogenesis-related (PR) proteins, which were initially identified in tobacco leaves responding hypersensitively to infection by tobacco mosaic virus (TMV) (Gianinazi et al., 1969, 1970; van Loon and van Strien 1999). Subsequent studies have identified PR proteins in a range of other species, including

Table 1 The families of PR-proteins

Family	Biological Properties
PR-1	Unknown
-2	β -1,3-glucanase
-3	Type I, II, IV, V, VI, VII chitinases
-4	Type I, II chitinases
-5	Thaumatin-like
-6	Proteinase inhibitor
-7	Endoproteinase
-8	Type III chitinase
-9	Peroxidase
-10	"Ribonuclease-like"
-11	Type I chitinase
-12	Defensin
-13	Thionin
-14	Lipid transfer protein
-15	Oxalate oxidase
-16	"Oxalate oxidase-like"
-17	"Proteinase-like"

Based on van Loon and van Strien (1999) and Christensen et al. (2002).

cereals and other major crops and have shown that their synthesis is also induced by infection with other microbes (viroids, bacteria, fungi) or treatment with chemical elicitors, notably salicylic acid or acetyl salicylic acid (aspirin). The latter is now known to relate to the role of salicylate in endogenous pathogen signalling pathways in the plant. The first PR proteins to be identified tended to be acidic and selectively extracted at low pH, highly resistant to proteolysis, and located in the intracellular space. However, it has since been shown that many PR proteins also exist in basic forms that are located in the vacuole.

Presently, 17 different classes of PR proteins have been defined, as summarized in Table 1. These clearly vary widely in their biological properties, but four types of activity are particularly noteworthy.

- (1) Enzyme inhibition, inhibitors of proteinases and α -amylases are widespread, but inhibitors of polygalacturonase (an enzyme secreted by plant pathogenic fungi) and other enzymes also occur.
- (2) Hydrolysis of polysaccharides or proteins, including lysozyme, which hydrolyzes bacterial cell walls, α -1,3-glucanases and endochitinases, which hydrolyze fungal hyphae and/or insect cuticles.
- (3) Destabilization of membranes resulting in leakage, including thionins, 2S albumins, thaumatin-related proteins lipid transfer proteins, and some defensins.
- (4) Binding to chitin (an α -1,4-linked polymer of *N*-acetylglucosamine present in fungal cell walls and insect cuticles), lectins, hevein, endochitinases, and some defensins.

However, the relationships between these activities and plant protection are still not fully understood, and many putative defensive proteins may have additional or different types of biological activity.

The most widespread type of protective protein is inhibitors of proteinases with over 20,000 studies having been published by 1991 (Richardson). Proteinase inhibitors are divided into twelve classes (see Shewry, 1999) defined on their amino acid sequences and target enzymes. Ten of these classes are active against serine proteinases (the most widespread class of proteinases), with the two remaining classes being active against carboxypeptidases and cysteine proteinases (the cystatins). There is little or no sequence homology between the different classes of proteinase inhibitors, which also vary considerably in their molecular masses and distribution in different groups of plants.

Seed Storage Proteins

These are classified into three groups based on their solubility properties. The globulin storage proteins are usually soluble in 0.5 to 1.0 M NaCl, although some types may be only sparingly soluble under these conditions. Globulins are classified into two groups on the basis of their sedimentation co-efficients (S_{20w}), determined by sedimentation ultracentrifugation. The literature on globulins is massive, and the reader is referred to review articles (Boulter and Croy, 1997; Derbyshire et al., 1976; Gatehouse et al., 1984; Shewry and Casey, 1999; Wright, 1987, 1988) for more detailed accounts and references to primary sources.

The 11S-12S globulins are the most widespread group and are present in monocotyledonous and dicotyledonous seeds, with homologues also present in gymnosperms (including conifers). They are particularly characteristic of legume seeds and, hence, are often called legumins. "Typical" legumins have M_r of about 300,000–450,000 and consist of six subunits of M_r about 60,000. These subunits are the products of a multigene family and are associated by noncovalent forces. Each subunit is post-translationally processed to give rise to acidic (M_r about 40,000) and basic (M_r about 20,000) chains, which are associated by a single disulphide bond. Thus, 11S/12S globulins give characteristic patterns of M_r 60,000 subunits when separated by SDS-PAGE without reduction and of M_r 40,000 + M_r 20,000 chains when reduced. The 11/12S globulins are rarely, if ever, glycosylated.

The 7/8S globulins also occur in a range of monocotyledonous and dicotyledonous species and are often called vicilins because of their presence in the *Viciae* group of legumes. They are typically trimeric proteins of M_r about 150,000–190,000, with subunit M_r ranging from about 40,000–80,000, but typically, about 50,000. Furthermore, proteolytic processing and glycosylation may occur, resulting in a diverse spectrum of components when separated by SDS-PAGE. For example, vicilin components ranging in M_r from about 12,000 to 50,000 are present in seeds of pea. Vicilins contain no disulphide bonds, and consequently, the spectrum of components revealed by SDS-PAGE is similar in the absence or presence of reducing agents.

Although the mature 11/12S globulins differ from the 7S globulins in being hexameric rather than trimeric, they are initially assembled in the endoplasmic reticulum of the seed into

trimers that are then transported to the vacuole, where proteolysis to release the acidic and basic polypeptides occurs, together with assembly into hexamers (Gatehouse et al., 1984; Müntz et al., 1993). Similarly the trimeric mature 7S globulins may undergo reversible aggregation into hexamers depending on ionic strength (Thanh and Shibasaki, 1979).

2S albumin storage proteins were first characterized in the early 1980s (Youle and Huang, 1981) and are now known to be a major family of storage proteins (Shewry and Pandya, 1999). They appear to be restricted to seeds of dicotyledonous plants, where they may accompany 11/12S globulins and/or 7/8S globulins. Most 2S albumins are synthesized as single chains of M_r 10,000–15,000 that are post-translationally processed to give small and large subunits that usually remain joined by disulphide bonds. As a consequence of their being rich in the sulphur-containing amino acid methionine, some 2S albumins have been a target for genetic manipulation of food crops to improve their amino acid balance. Thus, a 2S albumin from Brazil nut has been transferred into soy, with the result that extracts of the transgenic seed showed positive reactions for IgE binding and skin pricks when tested against patients suffering from allergy to Brazil nuts (Nordlee et al., 1996).

The prolamin storage proteins are characterized by their solubility in aqueous alcohols and insolubility in dilute salt solutions, either in the native state or after reduction of interchain disulphide bonds. Although all prolamins are rich in proline and glutamine, the combined proportions of these amino acids actually vary from about 30–70 mol% in different components. Similarly, their M_r range from about 10,000 to 100,000, and they include monomeric components, which either lack cysteine or have interchain disulphide bonds and high M_r (above 1×10^6) polymeric complexes stabilized by interchain disulphide bonds. Although it is not possible to define the structure and properties of a “typical” prolamin, almost all contain regions (domains), that either consist of repeated peptide motifs or are enriched in specific amino acids. These domains are responsible for the high proportions of glutamine, proline and other amino acids (e.g., methionine, histidine, glycine, phenylalanine) in specific groups of prolamins. These repetitive domains are usually accompanied by one or more nonrepetitive domains that contains most or all of the cysteine residues. The reader is referred to Shewry and Tatham (1990; 1999) and chapters in Shewry and Casey (1999) for further details and references.

The comparison of available amino acid sequences shows that all prolamins appear to be related, with exception of the major α -zeins of maize and homologues in other “panicoid” cereals (sorghum and millets). This suggests that they have evolved from a common ancestral protein or proteins, as discussed below.

Plant Protein Superfamilies

Since Victorian times, manual workers in the oil-seed (castor bean and cotton seed) milling and wheat milling and baking industries have become sensitized toward these seeds by inhalation

of air borne particles. Subsequently, the allergens in these seeds have been identified as either 2S albumins or α -amylase/trypsin inhibitors, leading to the view that these protein families should be considered to be potentially allergenic (Youle and Huang, 1981). The last two decades have seen a massive increase in our knowledge of plant proteins, with the availability of amino acid sequences from molecular cloning and of three-dimensional structures from x-ray crystallography and NMR spectroscopy. These allow proteins to be re-classified on the basis of their structural and evolutionary relationships, allowing the concept of allergenic plant protein families or pan-allergens to be studied more systematically. Most seed storage proteins form part of two larger protein superfamilies: the prolamin superfamily and the cupins. These superfamilies contain a number of major allergens that are thought to sensitize individuals via the gastrointestinal tract and will, therefore, be discussed in detail.

THE PROLAMIN SUPERFAMILY

This protein superfamily was initially identified on the basis of visual comparison of amino acid sequences, as described by Kreis et al. (1985). At this time, three groups of proteins were included, the prolamins of the Triticeae (barley, wheat, rye), cereal grain inhibitors of trypsin and α -amylase, and the 2S albumin storage proteins of dicotyledonous plants. Using computational methods of analysis, a comparison of protein sequences shows that the prolamin superfamily has six families: the prolamins of wheat, barley, oats, and rice: the minor prolamins of maize (β , γ and δ -zeins) and other panicoid cereals (sorghum, millets); the tryptophan-rich puroindolines of cereals; α -amylase inhibitors; 2S albumins and the non-specific lipid transfer proteins (ns LTPs) (Figure 1). It is of interest that all of these proteins, except the ns LTPs, appear to be seed-specific, at least in higher plants.

All members of the prolamin superfamily are characterized by a conserved pattern of cysteine residues that are located within

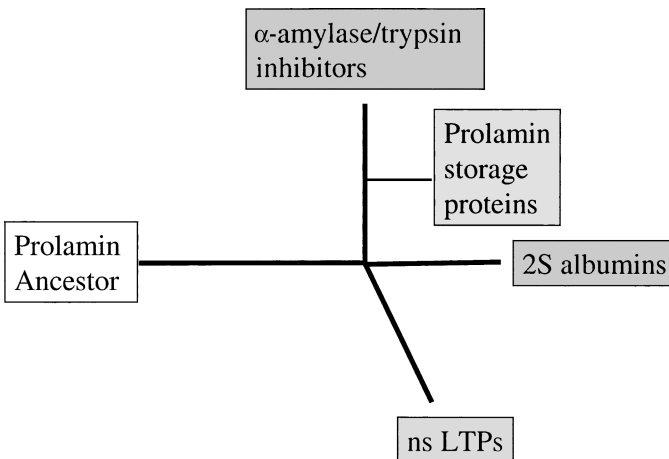


Figure 1 The evolutionary relationships of the main members of the prolamin superfamily.

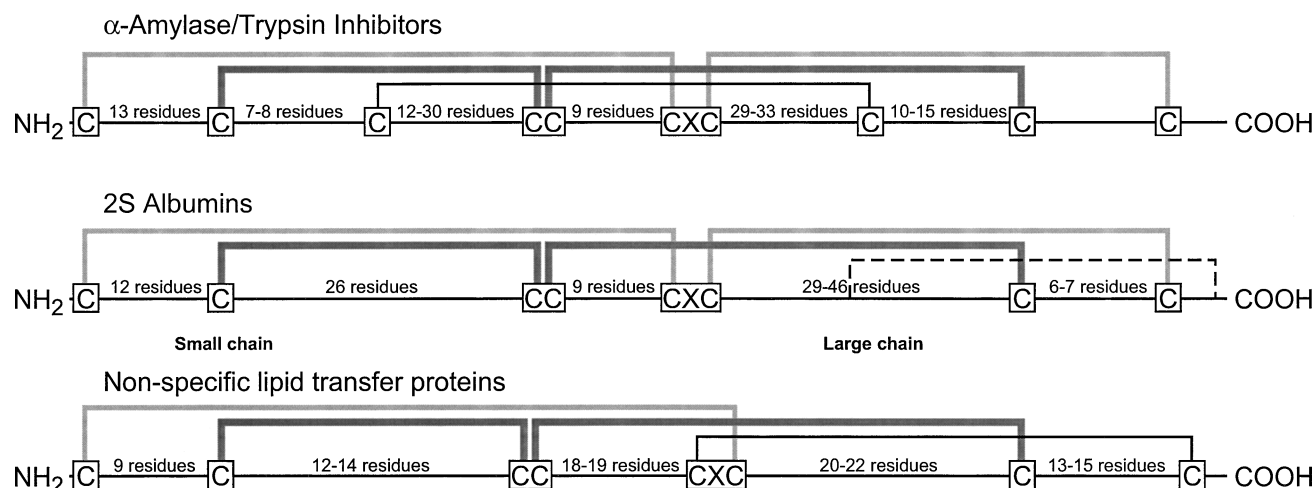
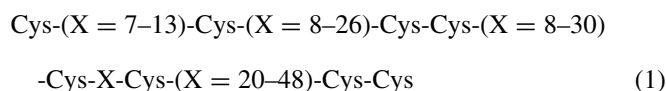


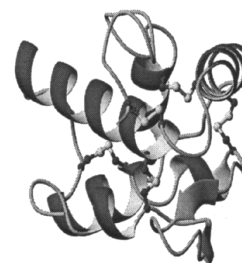
Figure 2 The conserved cysteine skeleton of the prolaminsuperfamily. Diagrammatic representation of the disulphide bonds formed between conserved cysteine residues. Ara h 6 and sesame 2S albumin have two additional cysteine residues compared with other members of the 2S albumin family. A putative disulphide bond is shown by the dotted line.

a sequence of about 100 amino acid residues (Figure 2). Either six or eight such cysteine residues are present. These residues form three or four intra-chain disulphide bonds. Most of the proteins contain characteristic Cys Cys and Cys X Cys motifs, where X represents any other residue and can be defined by the formula:

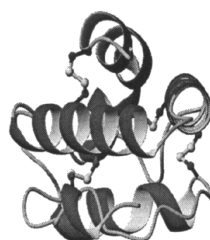


Most members of the prolaminsuperfamily also have low M_r (below about 15,000), with the conserved region accounting for almost the whole protein. However, certain prolaminsuperfamily proteins contain additional sequences that may be present as inserts or as separate protein domains. Particularly, most cereal prolamins storage proteins contain additional domains that consist of repeated amino acid sequences or are enriched in specific amino acid residues (e.g., methionine). The degree of sequence identity between the conserved regions of various prolaminsuperfamily proteins is low, but a comparison of known 3-dimensional structures of prolaminsuperfamily members demonstrates striking similarity at the structural level. Figure 3 illustrates this similarity by comparing the 3D structures of the ns LTP from maize, the α -amylase inhibitor from Indian finger millet (ragi), and the 2S albumin (napin) from oil seed rape. All the proteins have a related fold consisting of bundles of four α -helices stabilized by disulphide bonds using the structural classification of proteins (SCOP) data base (Murzin et al., 1995) as the "four α -helices folded leaf fold." Structural alignment shows that the positions of the α -helices are approximately conserved, but that there are some significant shifts. All the proteins also have four or five disulphide bridges with four distinguishable patterns of disulphides seen in the three-dimensional structures, and only two disulphides are conserved between superfamily members. While the pattern Cys-X-Cys is present in helix three of all

Wheat 0.19 Amylase Inhibitor



Maize n-LTP



Napin model 1

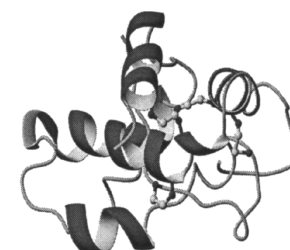
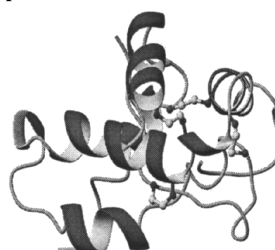


Figure 3 Aligned structures of members of the prolaminsuperfamily. A: A ribbon representation of maize non-specific lipid transfer protein (Shin et al., 1995). The conserved disulphides on the second α -helix are at the top right. B: A ribbon representation of wheat 0.19 α -amylase inhibitor (Oda et al., 1997). C: A ribbon representation of oilseed rape napin (Rico et al., 1996).

the superfamily members, these cysteines do not form the same disulphide bonds. This is most obvious for the ns-LTPs, where the order of disulphide bonds is different than that in the other members of the superfamily, and helix three has moved in space. The Cys-X-Cys pattern is also shifted by one turn of helix three in napin, compared with the α -amylase inhibitor. It is possible that this is a rare example of protein sequence being more highly conserved than protein structure.

The prolamin, ns LTP, α -amylase/trypsin inhibitor, and 2S albumin families all include proteins that are major food allergens. The characteristics of these allergens are summarized in Table 2 and discussed below, on a family-by-family basis.

Prolamins

Occurrence and Properties

Prolamins are restricted to the seeds, and particularly to the starchy endosperm tissues, of members of the grass family, where they form the major storage protein fraction in most species (Shewry and Tatham, 1999). They account, on average, for about 50% of the total grain proteins in most cultivated cereals (maize, wheat, barley, sorghum, millets), but only about 10% in oats and rice, where the major storage proteins are the 11/12S globulins. Prolamins are characterized by solubility in alcohol-water mixtures, either in the native state or after the reduction of inter-chain disulphide bonds (which stabilize polymeric complexes).

Structure and Sequence

Prolamins show immense variability in their structures, sequences, and properties. Two major groups can be recognized that do not appear to be related. The first is the α -type zeins of maize and related proteins in other tropical cereals, such as sorghums and millets. These proteins do not appear to belong to the prolamin superfamily or to any other recognized family of plant proteins. The second group comprises the minor zeins of maize (β , γ , δ -zeins), related proteins in other tropical cereals and all of the prolamins present in temperate cereals (wheat, barley, rye, oats), and rice. Since only wheat prolamins have been reported as allergens, only these will be discussed in detail.

Wheat prolamins correspond broadly to the gluten proteins and are classically divided into two groups: the monomeric gliadins, and polymeric glutenins (Figure 4). However, comparisons of amino acid sequences indicate that it is more valid to classify the proteins within these groups into three families. Thus, the monomeric α -type gliadins, γ -gliadins, and polymeric low molecular weight (LMW) subunits of glutenin are all sulphur-rich (S-rich) prolamins, while the monomeric ω -gliadins and the polymeric high molecular weight (HMW) subunits of glutenin are S-poor and HMW prolamins, respectively (Shewry et al., 1999).

The sequences of typical members of these groups are summarized in Figure 4 where it is shown that all have distinct

domain structures, including one domain consisting of repeated sequences. In the S-rich prolamins, a short unique *N*-terminal sequence is followed by a domain comprising repeated sequences based on one or two short motifs rich in proline (P) and glutamine (Q). Cysteine residues are present in the nonrepetitive *C*-terminal domain (and also close to the *N*-terminus in some LMW subunits). Sequence alignment shows that the characteristic Cys-Cys-(X = 9–19)-Cys-X-Cys motif of the prolamin superfamily is present in this domain, as shown for an α -gliadin as sequence in Figure 6A, and that it aligns more closely with the α -amylase/trypsin inhibitors than other members of the prolamin superfamily. Repeated sequences based on similar repeat motifs also comprise the bulk of the ω -gliadin as shown in Figure 4, with only short unique sequences at the protein *N*- and *C*-termini and no cysteine residues. Finally, the HMW subunits contain extensive repeated sequences based on two or three short motifs (in *y*-type and *x*-type subunits, respectively) that are not related to those present in the S-rich and S-poor prolamins. This repetitive domain is flanked by nonrepetitive domains that contain cysteine residues that are thought to form inter- and intrachain disulphide bonds, leading to the assembly of the subunits into high M_r polymers.

Allergens

Despite the widespread perception that cereals, particularly wheat are responsible for a large proportion of food allergies, there is a paucity of data regarding the role of cereal prolamins in allergic disease, although their role in the aetiology of coeliac disease is well defined. Studies on cereal allergies have been complicated by the low solubility of cereal prolamins in dilute salt solutions, routinely used in clinical studies, and the fact that cooked wheat flour provokes different symptoms compared with raw wheat flour (Guenard-Bilbault et al., 1999). Wheat has been shown to be involved in wheat-induced atopic dermatitis (Varjonen et al., 1995, 2000) and exercise-induced anaphylaxis (EIA) (Varjonen et al., 1997). EIA is a reaction that certain patients experience only when exercising after the consumption of a problem food. Two allergens have been described: a γ -gliadin and an α -gliadin (Palosuo et al., 1999). Others (Takizawa et al., 2001; Tanabe et al., 1996) have also found the prolamins to be major cereal allergens and in one instance LMW subunits of glutenin containing the pentapeptide motif Gln-Gln-Gln-Pro-Pro have been identified as the major protein binding to IgE from patients with atopic dermatitis to wheat (Tanabe et al., 1996). Maruyama et al. (1998) have also found that a γ - and an α -gliadin were also IgE-reactive prolamins.

Non-Specific Lipid Transfer Proteins

Occurrence and Properties

The ns LTP family differs from other members of the prolamin superfamily, in that they are not restricted to seed tissues. They were initially defined on the basis of their ability to transfer

Table 2 A summary of allergenic plant food proteins

Allergen family	Function in plant	Biochemical characteristics	Botanical source	Allergen name	MWt	Sequence accession	Reference
Prolamin-superfamily							
Prolamins	Seed storage protein	Soluble in aqueous mixtures of alcohols and reducing agents. They are rich in proline and glutamine and possess a characteristic central domain comprising tetra- and hexapeptide repeats. Only certain members of this multigene family are allergens.	Wheat (<i>Triticum aestivum</i>)	α - and γ -gliadin LMW glutenin	~31–42.00 ~33–44.00	BAA12318 BAA11251 BAA23162	Paluoso et al., 1999; Varjonec et al., 1995, 1997, 2000; Takizawa et al., 1996; Maruyama et al., 1998.
Non-specific lipid transfer proteins (ns LTPs)	Function uncertain: maybe involved in transport of suberin monomers.	Members of the ns LTP family are thermostable, generally only becoming unfolded at temperatures above 90°C. These proteins are also resistant to digestion and are not glycosylated in plants.	Peach (<i>Prunus persica</i>) Apple (<i>Malus domestica</i>) Maize (<i>Zea mays</i>)	Pru p 1 Mal d 3 Zea m 14	9,178 11,410 9,054	P81402 Q9M5X7 P19656	Pastorello et al., 1999a,b Anchez-Monge et al., 1999; Pastorello et al., 1999c. Pastorello et al., 2000b. Pastorello et al., 2000a. Scheurer et al., 2001 James et al., 1997 Curionie et al., 1991 Pastorello et al., 2000b.
α -Amylase/trypsin inhibitors	Provide protection against degradative proteases and amylases produced by insect pests and pathogens	Soluble in water, dilute salt and chloroform/methanol water mixtures. They exist as monomeric, dimeric and tetrameric forms. They appear to be thermostable, retaining their allergenicity following processes such as brewing.	Apricot (<i>Prunus armeniaca</i>) Cherry (<i>Prunus avium</i>) Wheat (<i>Triticum aestivum</i>) Barley (<i>Hordeum vulgare</i>) Maize (<i>Zea mays</i>) Rice (<i>Oryza sativa</i>)	Pru ar 3 Pru av 3 Hor v 1 Rag 1, 2, 5, 5.b 14, 14b, 16, 17	9,178 9,200 15,000 16,000 16,000 14–16,000	G7404406 AAF26449 CAA45085 Q01881 S59922 S59924 S59925	2000b. Pastorello et al., 2000a. Scheurer et al., 2001 James et al., 1997 Curionie et al., 1991 Pastorello et al., 2000b. Nakase et al., 1996; Alvarez et al., 1995; Izumi et al., 1992
2S Albumins	Seed storage protein	Synthesised as a single chain and generally cleaved to give a conserved 3,000 dalton polypeptide disulphide linked to a slightly more variable 9,000 dalton polypeptide. Unusually, sunflower has a single chain albumin, which forms dimers. Allergen present in raw and roasted seeds.	Walnut (<i>Juglans regia</i>) Yellow mustard (<i>Sinapis alba</i>) Oriental mustard (<i>Brassica juncea</i>) Brazil nut (<i>Bertholletia excelsa</i>) Peanut (<i>Arachis hypogaea</i>) Sesame (<i>Sesamum indicum</i>) Sunflower (<i>Helianthus annuus</i>)	Jug r 1 Sin a 1 Bra j 1 Ber e 1 Ara h 2 Ara h 6 Ara h 7 Ses i 1 SFA-8	14,000 14,100 14,662 12,000 17,500 14,500 15,800 ~9,000 12,147	JRU66866 P15322 P80207 CAA38362 L77197 AF091737 AF092846 AAD42943 12,000	Teuber et al., 1998 Mendez-Arias et al., 1998 Monsalve et al., 1993 Pastorello et al., 1998 Burks et al., 1992a; Stanley et al., 1997; Kleber-Janick et al., 1999 Pastorello et al., 2001 Kelly et al., 2000

Cupin superfamily 7S (vicilin-like) Globulins	Seed storage protein	Proteins comprise 50–60,000 dalton subunits assembled into 180,000 dalton trimers. The trimers may be glycosylated, having a single N-linked glycosylation site. In some plant species, such as pea, they are subject to extensive posttranslational proteolysis and in others, such as soybean, the trimers comprise a mixture of subunits. The trimer is resistant to pepsinolysis. Proteins are thermostable up to 75°C, forming large aggregates above that temperature.	Peanut (<i>Arachis hypogaea</i>) Walnut (<i>Juglans regia</i>) Lentils (<i>Lens culinaris</i>) Soya (<i>Glycine max</i>)	Con-arachin, Ara h 1 Jug r 2 β -Conglycinin	60,000 47,000 α : 60,000 β : 50,000	L34402 AAB41308 α : g13916 α' : P11827 β : P25974	Burks et al., 1991, 1992b Teuber et al., 1999 Sanchez-Monge et al., 2000 Burks et al., 1998
11S (legumin-like) Globulins	Seed storage protein	Comprise an acidic polypeptide disulphide bonded to a basic polypeptide. Subunits assembled into 360,000 dalton hexamers. Are products up of a multigene family and thermostable up to 92-5°C. Forms intermediates on trypsinolysis which retain much of the quaternary structure. Only rarely glycosylated.	Peanut (<i>Arachis hypogaea</i>) Soy (<i>Glycine max</i>)	Arachin, Ara h 3, Ara h 4, Glycinin	61,000 54,000	AF093541 AF086821 Gy1 (A1aBx): P04776 Gy2 (A2B1a): P04405 Gy3 (AB): P11828 Gy4 (A4/5B3): P02858 Gy (A3B4): P04347	Burks et al., 1992b; Rabjohn et al., 1999 Beardslee et al., 2000; Nielsen et al., 1989.
Cysteine protease C1 family	Cysteine proteases	Have conserved Gln, Cys, His and Asn residues at the active site and whilst sharing sequence homology may not have protease activity. Synthesised as a precursor protein, some are N-glycosylated.	Kiwi fruit (<i>Actinidia chinensis</i>) Soy (<i>Glycine max</i>)	Actinidin Glym Bd 30K; P34; Gly m 1	30,000 34,000	Q43367 P22895	Pastorello et al., 1998a. Ogawa et al. (1993)

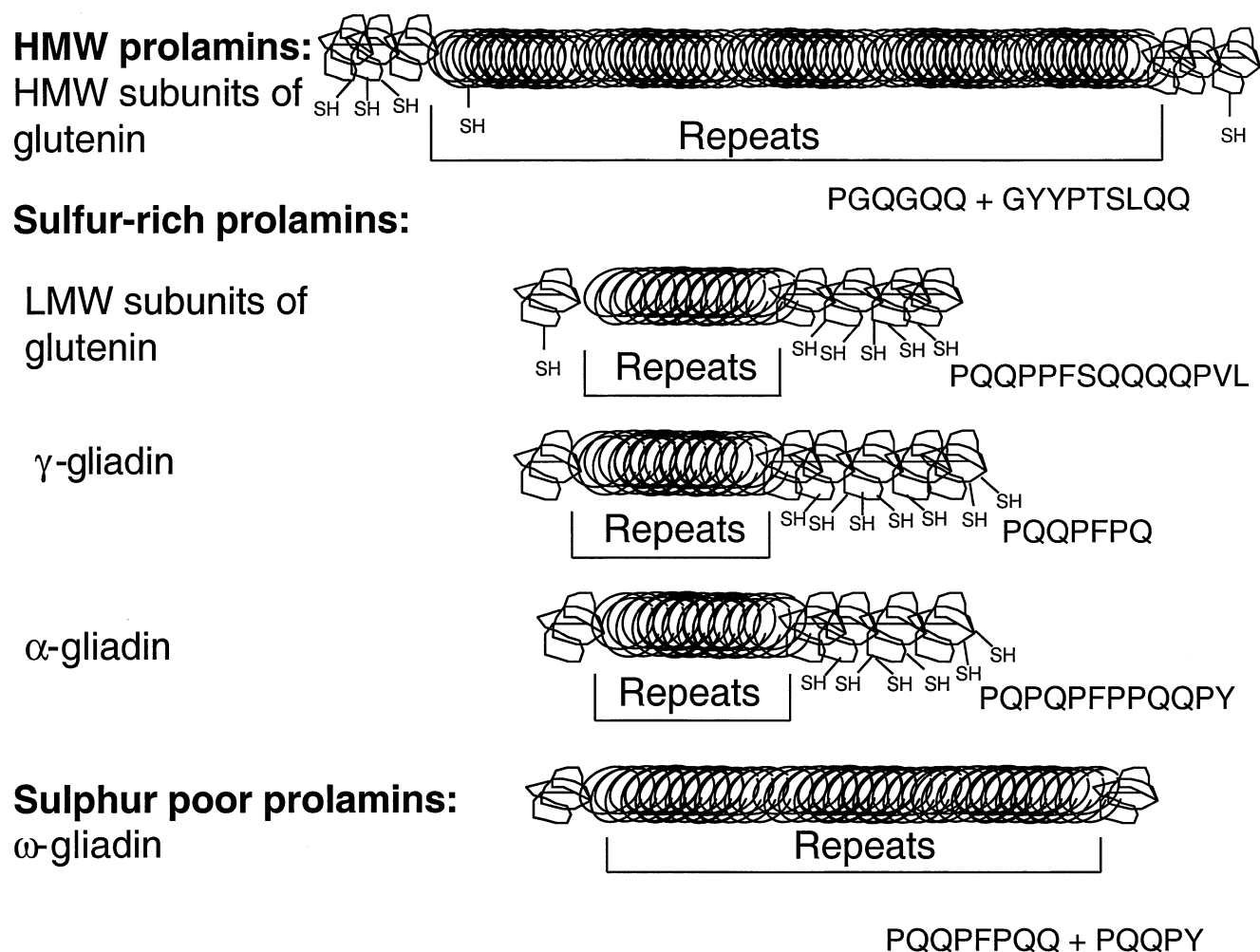


Figure 4 Summary of the domain structures of “typical” wheat prolamins. The consensus amino acid sequences of the repeat motifs are shown using standard single letter abbreviations: G, glycine; F, phenylalanine; L, leucine; P, proline; Q, glutamine; S, serine; T, threonine; V, valine; Y, tyrosine. The figure is based on reported amino acid sequences (Anderson et al., 1989; Bartels et al., 1986; Colot et al., 1987; Hsia and Anderson, 2001; Kasarda et al., 1984).

phospholipids from liposomes to mitochondria *in vitro*, but this activity lacks specificity in terms of the type of phospholipid and membrane. Their *in vivo* function remains unknown, but the observations that ns LTPs frequently occur in epidermal tissues and that components from radish and onion seeds inhibit fungal growth *in vitro* indicate that they may be protective. It has also been suggested that they may be involved in the transport of cutin and suberin monomers to the site of waxy cuticle synthesis. There is also a relationship between ns LTP expression and cuticle wax thickness (Douliez et al., 2000). Such a role is compatible with the involvement of ns LTPs in plant defence responses, and some ns LTPs have been designated as belonging to PR (pathogen response) group 14 (van Loon and van Strien, 1999).

Structure and Sequence

More than 100 plant ns LTPs have been sequenced, and the aligned sequences of a number from various plant seeds and tissues are shown in Figure 5. Comparisons of sequences show that ns LTPs are closely homologous, comprising 67–139 residues

($M_r \sim 7,000$ to at least 15,000) with eight conserved cysteines that form four intrachain disulphide bonds. However, some proteins with ns LTP activity have sequences that are much more distantly related, such as the ns LTP of castor bean (Tsuboi et al., 1991).

The ns LTPs are the best characterized members of the prolamins superfamily at the structural level, with 3D-structures available of ns LTPs from maize (Gomar et al., 1988; Gomar et al., 1996; Shin et al., 1995), rice (Lee et al., 1988; Poznanski et al., 1999), barley (Heinemann et al., 1996), and wheat, the latter determined both by X-ray and NMR methods (Charvolin et al., 1999; Gincel et al., 1994; Tassin-Moindrot et al., 2000). There is clear homology between these ns LTPs, with 52 out of 90 structurally aligned residues being identical (57.8%) between maize and wheat, 46 out of 85 (54.1%) between maize and barley, and 70 out of 89 (78.7%) between maize and rice. The structures of some ns LTPs (including those from barley and wheat) have been determined for the free protein and when complexed with one or two lipids. The free and lipid-bound forms of ns LTPs show structural differences, most notably in the

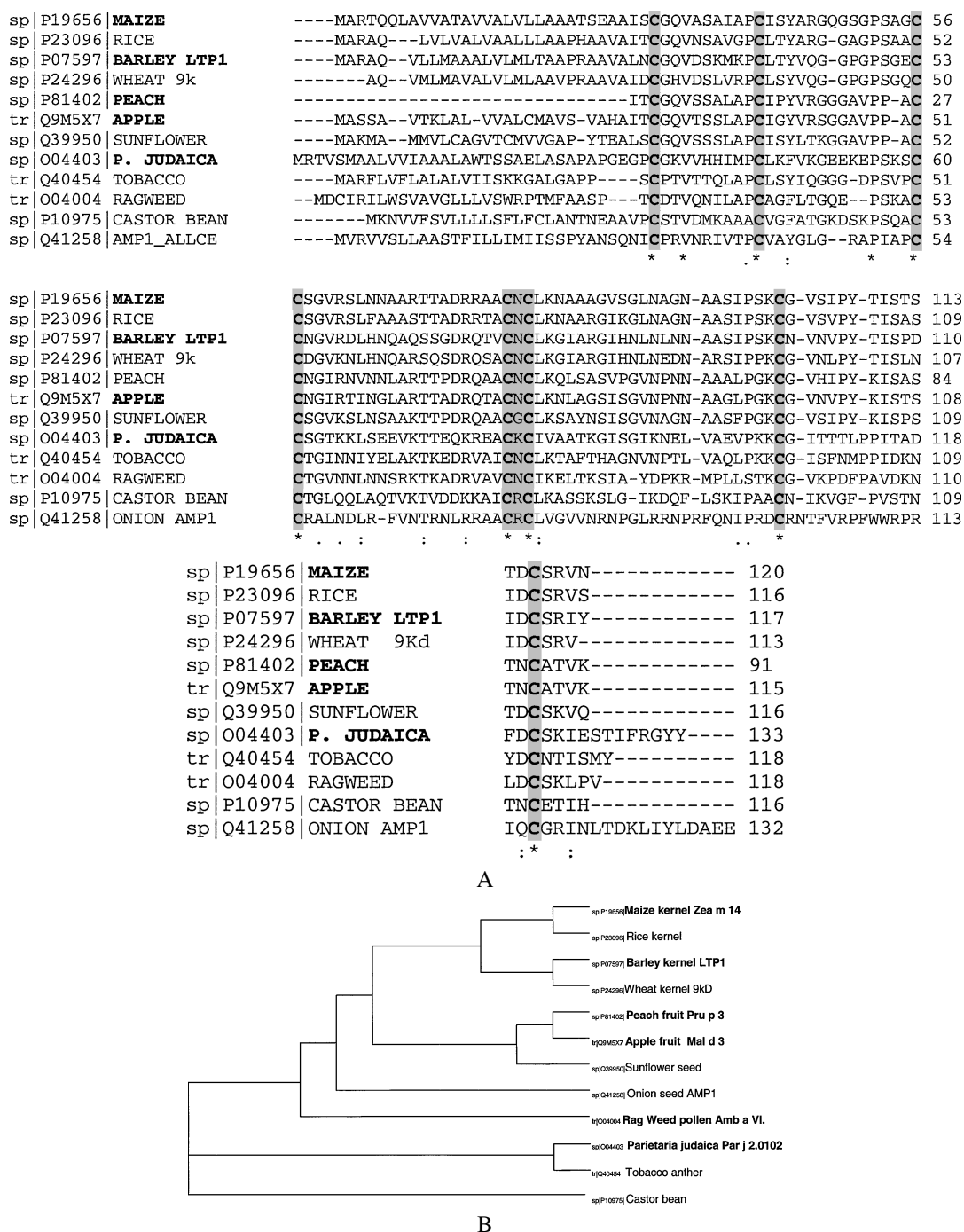


Figure 5 Aligned sequences (A) and dendrogram (B) for a selection of allergenic and non-allergenic plant ns LTPs. Allergens are shown in bold. *, conserved residues; . conserved substitutions; : semi-conserved substitutions. Conserved cysteine residues are shaded. Allergen names are preceded by database accessions: sp—Swiss Prot; tr—TREMBL.

C-terminal residues 75–84 of the wheat ns LTP, although smaller changes are observed in the same region of the maize ns LTP on binding lipid.

Allergens

The peach ns LTP allergen was the first to be characterized (Pastorello et al., 1999a; Sanchez-Monge et al., 1999), but ns

LTPs have since been found to be major allergens in various foods, leading to the family termed as ‘pan-allergens’ (Sanchez-Monge et al., 1999). The high degree of sequence homology shown by the members of this family results in the ns LTPs being highly cross-reactive allergens. Thus, in a study of maize allergy, the IgE from a significant proportion of the patients recognized both maize LTP and LTPs from a variety of fruits, as determined by immunoblotting (Pastorello et al., 2000b). LTP

allergens have also been identified in pollen, where patients had pollen allergies without associated food allergies (Duro et al., 1996; Tejera et al., 1999) and have also been identified in an individual with an occupational allergy to spelt, who had symptoms on consuming a variety of fruits (Pastorello et al., 2001a). Thus, ns LTPs appear to behave in a similar fashion to the 2S albumins, where sensitization by inhalation and/or the oral route can occur.

Ns LTPs are remarkably thermostable, retaining their allergenic activity in processed foods, such as sterilized peach juice (Brenna et al., 2000), beer (Asero et al., 2001; Curioni et al., 1999), and fermented products, such as wine (Garcia-Robiana et al., 2001). Thus, the ns LTP of barley only begins to unfold following heating above 95°C (Lindorff-Larsen and Winther, 2001) and can survive the brewing process in a modified but immunoreactive form (Bech et al., 1995). The proteins are also resistant to proteolysis by a range of enzymes including pepsin and trypsin (Asero et al., 2000; Lindorff-Larsen and Winther, 2001), despite having a number of internal tryptic cleavage sites.

While not especially abundant in fruits or seeds, ns LTPs can become concentrated in epidermal tissues and associated with the waxy layers (Douliez et al., 2000). Thus, one strategy that has been employed successfully to reduce the allergen content of peach juice is to use chemical peeling prior to juicing to remove the bulk of the peel associated ns LTP allergens (Brenna et al., 2000).

Trypsin and α -Amylase Inhibitors

Distribution and Properties

Seeds of the triticeae (barley, wheat, and rye) and some other cereals (notably maize, rice, and ragi) contain a family of inhibitors of trypsin and exogenous (i.e., non-plant) α -amylases. These were initially identified in extracts made with chloroform/water mixtures (called CM proteins), but are also present in extracts made with water, dilute saline, or alcohol/water mixtures. More detailed studies have revealed a range of monomeric, dimeric, and tetrameric forms, with the individual subunits being either inactive or inhibitory to trypsin (and sometimes other proteinases), α -amylases from insects (including pests), or both enzymes (i.e., the inhibitors are bifunctional).

Structure and Sequence

A number of cDNAs and/or genes for α -class cereal inhibitor subunits have been isolated, showing that the encoded proteins comprise between 121 and 171 amino acid residues before removal of the signal peptides. Thus, the inhibitors tend to be significantly larger than the ns-LTPs, with M_r ranging from about 12,000 to 16,000 (Figure 6). Eight, nine, or ten cysteine residues are present that appear to form four or five interchain disulphide bonds. Although some subunits contain an unpaired cysteine

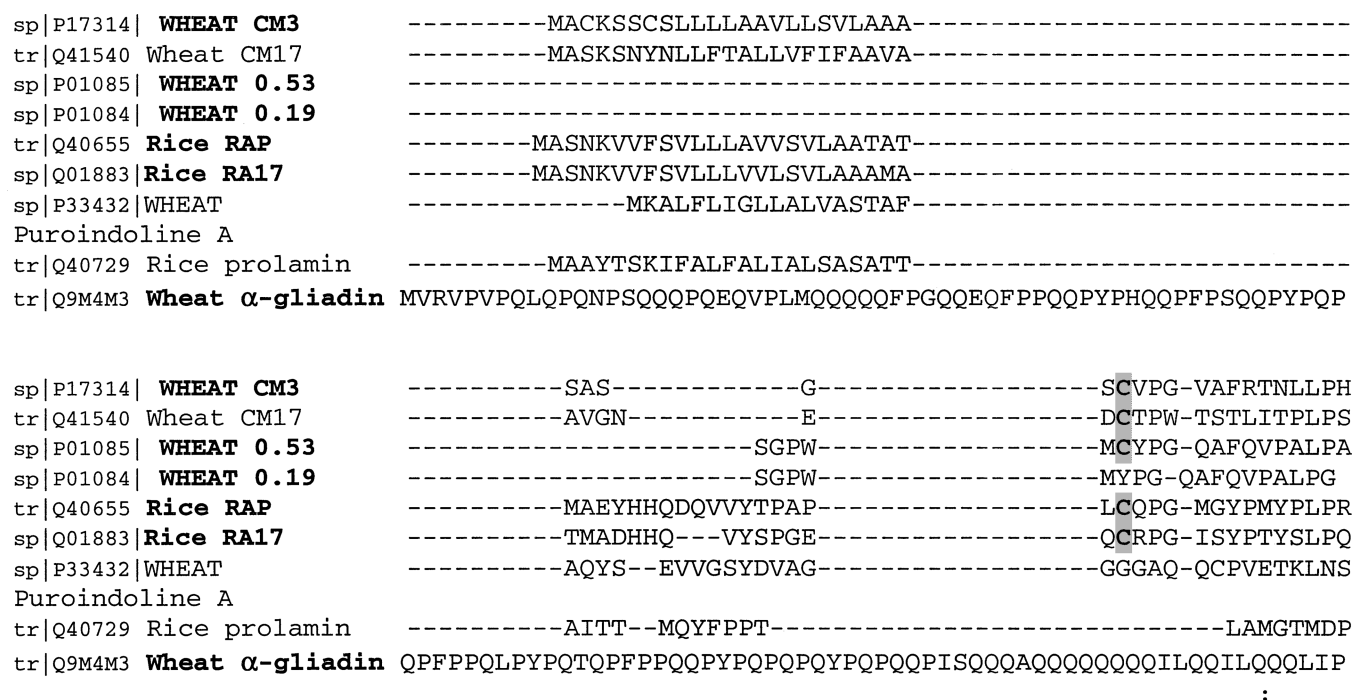


Figure 6 Aligned sequences (A) and dendrogram (B) for a selection of allergenic and non-allergenic α -amylase inhibitors. Allergens are shown in bold. *, conserved residues; ., conserved substitutions; :, semi-conserved substitutions. Conserved cysteines are shaded. Allergen names are preceded by database accessions: sp—Swiss Prot; tr—TREMBL. (Continued)

```

sp|P17314| WHEAT CM3          CRDYVLQQTCTGFTPGSKLPEWMT-----SASIYSPGKPYLAKLYCCQELA
tr|Q41540| Wheat CM17          CRNYVEEQACRIEMPG-----PPYLAKQECCEQLA
sp|P01085| WHEAT 0.53        CRPLLRLLQCNG---SQ-----VPEAVLRDCCQLA
sp|P01084| WHEAT 0.19        CRPLLRLLQCNG---SQ-----VPEAVLRDCCQLA
tr|Q40655| Rice RAP          CRALVKRQCVGRGTAA-----AAEQVRRDCCRQLA
sp|Q01883| Rice RA17        CRTLVRRQCVGRGASA-----ADEQVWQDCCRQLA
sp|P33432| WHEAT              CRNYLLDRCSTMKDFPVTW-----RWWKWWKGGCQELLGECSSRLG
Puroindoline A
tr|Q40729| Rice prolamin      CRQYMMQTLGMSSTAMF-----MSQPMALLLQCCMQLO
tr|Q9M4M3| Wheat α-gliadin   CRDVLVLLQHNIHAH-----SSQVLQSSYQQLQQLCCQQLF
*           :                               **  .:

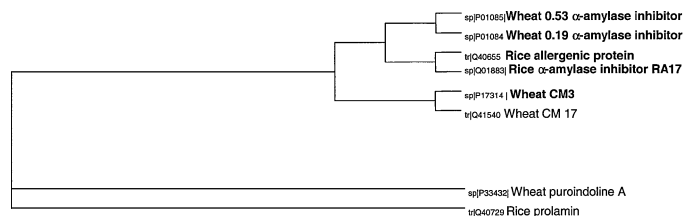
sp|P17314| WHEAT CM3          EISQ--QCRCEALRYFIALPVPSQPVDPR-----
tr|Q41540| Wheat CM17          NIPQ--QCRCQALRYFMGPK-----
sp|P01085| WHEAT 0.53        HISE--WCRCGALYSMLDSMYKEHGAQE-----
sp|P01084| WHEAT 0.19        DISE--WPRCGALYSMLDSMYKEHGVSE-----
tr|Q40655| Rice RAP          AVDD-SWCRCEAISHMLGGIYRELGAPDV-----
sp|Q01883| Rice RA17        AVDD-GWCRCGALDHMLSGIYRELGATEA-----
sp|P33432| WHEAT              QMPP--QCRCNIIQGSIQGDLGGIFGFQR-----
Puroindoline A
tr|Q40729| Rice prolamin      GMMP--QCHCGTSCQMMQSMQVICA--G-----
tr|Q9M4M3| Wheat α-gliadin   QIPE--QSRCQAIHNVVHAILLHHHQQQQQPSSQVSYQQPQEYPSGQGSFQSSQQ---
*

sp|P17314| WHEAT CM3          -----SGNVGESGLIDLPGCPREMOWDFVRLLVAPGQCNLATIHNVRYCPAVEQPLWI-
tr|Q41540| Wheat CM17          -----SRPDQSGLMELPGCPREVQMNFPILVTPGYCNLTTHVNTTPYCLGMEESQWS-
sp|P01085| WHEAT 0.53        -----GQAGTGAFPRCRREVVK--LTAASITAVCRLPIVVDASGDGAYVCKDVAA
sp|P01084| WHEAT 0.19        -----GQAGTGAFPRCRREVVK--LTAASITAVCRLPIVVDASGDGAYVCKDVAA
tr|Q40655| Rice RAP          -----GHPMSEVFRGCRRGDLE--RAAASLP AFCNVDI PNGGGGV CYWLARSGY-
sp|Q01883| Rice RA17        -----GHPMAEVFPGCRRGDLE--RAAASLP AFCNVDI PNGPGGV CYWLGYPRTP
sp|P33432| WHEAT              -----DRASKVIQEAKNLPPRCNQGPNCNIPGTIGYYW-----
Puroindoline A
tr|Q40729| Rice prolamin      -----LGQQQMMKMAMQMPYMCNMAPVNFQLSSCGCC-----
tr|Q9M4M3| Wheat α-gliadin   -----NPQAQGSVQPQQLPQFQEI RN LALQTL PAMCNVYIPPYCSTTIAPFGIFGTN
*

sp|P17314| WHEAT CM3          ----
tr|Q41540| Wheat CM17          ----
sp|P01085| WHEAT 0.53        YPDA
sp|P01084| WHEAT 0.19        YPDA
tr|Q40655| Rice RAP          ----
sp|Q01883| Rice RA17        RTGH
sp|P33432| WHEAT              ----
Puroindoline A
tr|Q40729| Rice prolamin      ----
tr|Q9M4M3| Wheat α-gliadin   YR--

```

A



B

Figure 6 (Continued).

residue, this does not appear to form an intersubunit disulphide bond, the oligomeric forms of the inhibitors are stabilized only by noncovalent forces.

Members of the inhibitor family for which structural information is available include those from wheat (Oda et al., 1997), maize (Behnke et al., 1988), and ragi (Strobel et al., 1995; Gourinath et al., 2000). Their three-dimensional structures are more divergent than those of the ns-LTP family. The maize and ragi inhibitors are quite closely related, with 71 identities in 112 structurally aligned residues (63.4%), and the homology may be higher because some identical residues are displaced by loops moving within the molecules. However, the wheat and maize inhibitors share only 25 identities out of 82 structurally aligned residues (30.5%). Four disulphide bridges are conserved within the family, a fifth disulphide bridge is present in the inhibitors from wheat and ragi, but not in that from maize.

Allergens

Allergens belonging to the cereal inhibitor family are of particular interest, because they are involved in both respiratory and food allergies. Although the former are strictly outside the scope of this study, it is nevertheless of interest to consider their structures and properties in relation to those of related non-allergenic proteins and dietary allergens.

Bakers' asthma is an occupational allergy developed by millers, bakers, and others who handle cereal flours. Most subunits of the α -amylase/trypsin inhibitors present in seeds of barley, wheat, and rye react with IgE from allergic patients when tested *in vitro*. However, the relative activities of the different subunits vary greatly, with the most active being glycosylated forms of subunits present as either monomeric and tetrameric α -amylase inhibitors in barley, or as tetrameric α -amylase inhibitors in wheat. These subunits are designated BMAI-1, BTAI-CMb, and WTAI-CM16, respectively (see Carbonero and García-Olmedo, 1999). The same allergens have been characterized in orally sensitized wheat allergic individuals, although only one $M_r \sim 15,000$ subunit was involved (James et al., 1997). An M_r 16,000 beer allergen has also been described that originates from barley and appears to belong to the α -amylase inhibitor family (Curioni et al., 1999). Similarly, a major M_r 16,000 allergen in maize probably also belongs to this family (Pastorello et al., 2000b).

The other well characterized food allergy to the cereal inhibitor family is in rice. The major rice allergens are inhibitors of human salivary α -amylase with M_r of about 14,000–16,000 (Nakase et al., 1996). More than ten different cDNA clones have been isolated indicating the presence of a multigene family (Alvarez et al., 1995; Izumi et al., 1992; Nakamura and Matsuda, 1996). However, antisense expression of a single sequence has resulted in decreases in the total allergen content of transgenic rice seed from about 300 $\mu\text{g}/\text{seed}$ (20 mg) to 60–70 $\mu\text{g}/\text{seed}$, which was associated with a decrease in transcripts to about 20% of the wild type levels.

2S Albumins

Distribution and Properties

2S albumin storage proteins have been characterized, and in some cases, corresponding cDNAs and/or genes isolated, from a botanically diverse range of dicotyledonous plants. They include major temperate crops, such as brassicas (oilseed rape, mustards), sunflower and legumes (lupin, pea), and tropical crops imported into Western Europe and Northern America, such as Brazil nut, sesame, and castor bean. An interesting feature is the presence in some species of methionine-rich albumins. These are present, together with methionine-poor forms, in sunflower, but appear to be the only type of 2S albumin present in Brazil nut (Shewry and Pandya, 1999).

Several types of biological activity have been reported for 2S albumins, including inhibition of proteinases (trypsin, subtilisin, and α -chymotrypsin) in some brassica species (Svendsen et al., 1989, 1994; Genov et al., 1997) and antifungal activity by membrane permeabilization in radish (*Raphanus sativus*) (Terras et al., 1992, 1993). Furthermore, the related mabinlin II protein in the seeds of the tropical plant mabinlin (*Capparis masaiikai*) is one of a small number of naturally occurring sweet proteins (Liu et al., 1993; Nirasawa et al., 1993). Activity as an antagonist of calmodulin (Polya et al., 1993) has also been proposed.

Sequences and Structures

Most 2S albumins have a "typical" structure comprising small and large "subunits" and associated by four interchain disulphide bonds, as previously discussed. However, variant forms do exist. For example, sunflower albumins are not post-translationally processed to give two subunits, while the subunits of pea albumin, which appears to belong to a different structural family, are not linked by disulphide bonds (Figure 7). Furthermore, in sunflower and castor bean, two mature albumins may be encoded by a single mRNA and synthesized as a single protein that is post-translationally cleaved to give two single subunits (sunflower) or two heterodimeric proteins (castor bean).

As with other groups of storage proteins, the 2S albumins show a high level of sequence variation (Figure 8), although they share the conserved disulphide structure common to all members of the prolamin superfamily. For example, the napin precursor has only 22% identical aligned residues with the Brazil nut 2S albumin, although restriction of the sequence comparison to residues 38–138 of the Brazil nut 2S against the two mature napin chains raises this to 31%. Thus, while the 2S albumins have a common architecture, their sequences are sufficiently divergent, making it impossible to correctly align every residue.

The 2S seed storage proteins are represented by a single NMR structure of the napin protein (Rico et al., 1996) and are more closely related in structure to the amylase inhibitor family than any other member of the prolamin superfamily, sharing the same

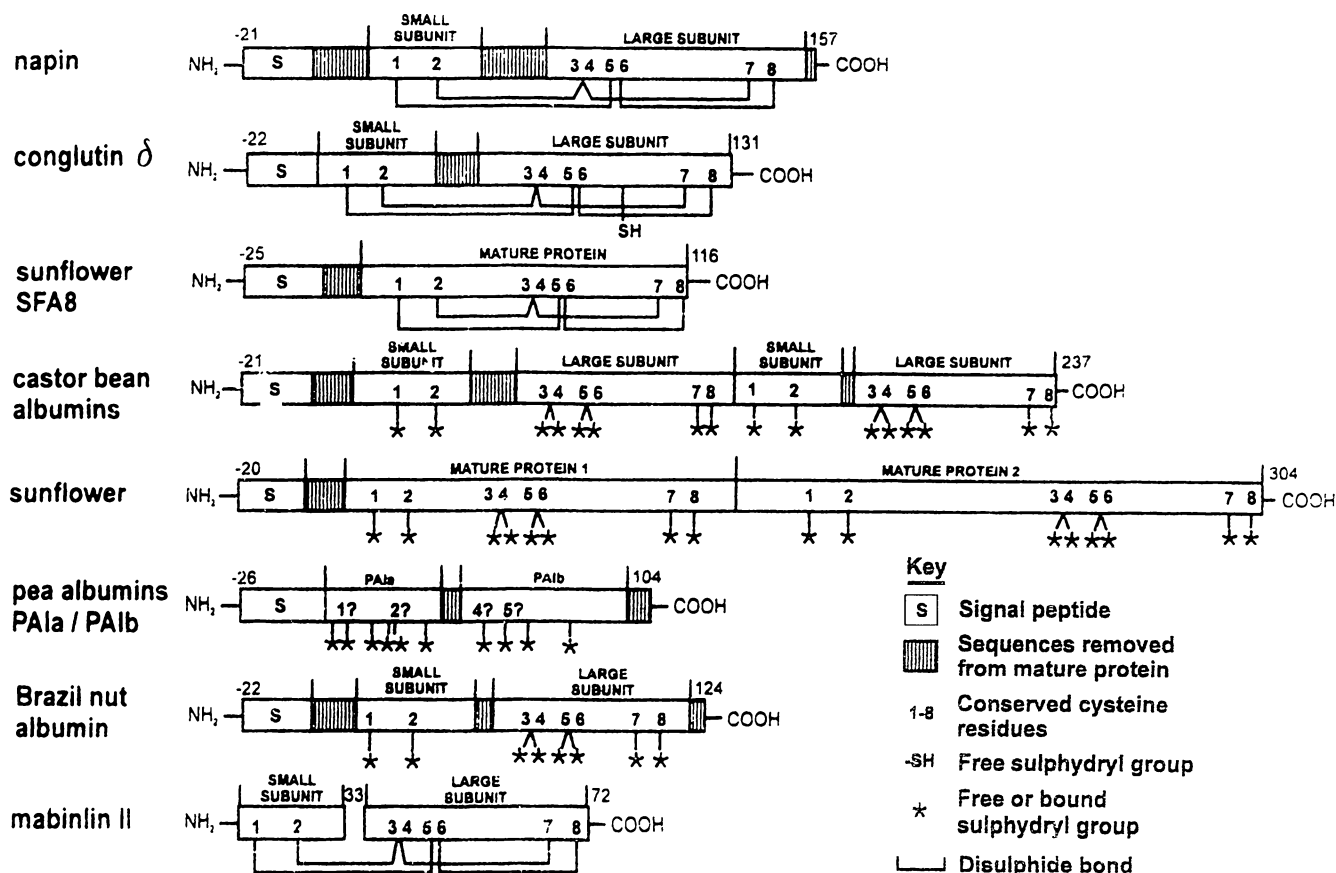


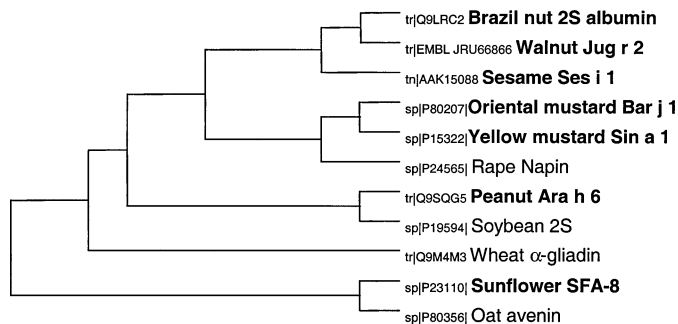
Figure 7 Schematic diagrams of various types of 2S albumins, indicating their origins from precursor proteins and their disulphide structures. Conserved cysteine residues are numbered 1–8 and free sulphhydryl groups shown as—SH. Cysteine residues whose behaviour is unknown are indicated by *. The precise correspondence between the cysteines in the pea albumins, PA1a and PA1b, and those in the other albumins is not known and potentially conserved residues are indicated by the number and ? The origin of the two mabinlin II subunits is not known, but they are presumed to derive from a single precursor protein as are the other heterodimeric albumins. The albumins, PA1a and PA1b differ from the dimeric albumins, in that the two subunits do not remain associated by inter-chain disulphide bonds. These are based on data and alignments in Crouch et al. (1983), Ericson et al. (1986), Lilley and Inglis (1986), Higgins et al. (1986), Irwin et al. (1990), Kortt et al. (1991), Egorov et al. (1996), Thoits et al. (1996) and taken from Shewry and Pandya (1979) with permission.

tr	Q9LRC2	Brazil nut	-----MAKISVA-AAALLVLMALGHATAF-----
tr	EMBL_JRU66866	walnut	-----AALLVALLFVANAAAF-----
tn	AAK15088	Sesame	-----MAKKLALAALLVAMVALASATTY-----
sp	P24565	Napin	-----
sp	P80207	Orient Must	-----
sp	P15322	Yellow Must	-----
tr	Q9SQG5	Ara h 6	-----AHASA-----
sp	P19594	Soybean	-----MTKFTILLISLLFCIAHTCS-----
sp	P23110	SFA-8	-----MARFSIVFAAGVLLLVAMA-----
sp	P80356	Avenin	-----MKTFLIFALLAMAATMATA-----
tr	Q9LRC2	Brazil nut	-----R-----ATVTTTVVEEE-----NQEECREQM-----RQQLSH
tr	EMBL_JRU66866	walnut	-----R-----TTITMEIDEDIDNPRRRGEGCREQIQ-----RQQLNH
tn	AAK15088	Sesame	-----T-----TTVTTTAIDDE-----ANQSSQCRQQLQ-----GRQ-FRS
sp	P24565	Napin	-----QPQKQREFQ-----QEQLRA
sp	P80207	Orient Must	-----AGPFRFPKCRKEFQ-----QAQLRA
sp	P15322	Yellow Must	-----PAGPFRIKCRKEFQ-----QAQLRA
tr	Q9SQG5	Ara h 6	-----M-----RRERG-----RQGDSSSCERQVD-----GV-NLKP
sp	P19594	Soybean 2S	-----A-----SKW-----QHQQDSCRQQLQ-----GV-NLTP
sp	P23110	SFA-8	-----PVSEASTTTITTTIIEENPYGRGRTEGSCYQOME-----EAEMLNH
sp	P80356	Avenin	-----QFDPSEQYQPYPEQQQPILQQQMLLQQQQMLLQQQPL-LQVLQQQLNP

Figure 8 Aligned sequences (A) and dendrogram (B) for a selection of allergenic and non-allergenic 2S albumins. Allergens are shown in bold. * conserved residues; . conserved substitutions; : semi-conserved substitutions. Conserved cysteines are shaded. Allergen names are preceded by database accessions: sp—Swiss Prot; tr—TREMBL. (Continued on next page).

tr Q9LRC2	Brazil nut		CRMYMRQQMEESPYQTM-----PRRGMEPHMSECCCEQLE
tr EMBL_JRU66866			
	walnut		CQYYLRQQSRSGGYDED-----NQRQHFRQCCQQLS
tn AAK15088	Sesame		CQRYLSQGRSPYGGEEDEVLE-----MSTGNQQSEQSLRDCQQQLR
sp P24565	Napin		CQQWIRQQLAGS---PF-----QSGPQEGPWLRQCCNELY
sp P80207	Orient Must		CQQWLHKQAMQSGSGPQ-----PQGPPQRPPLLQCCNELH
sp P15322	Yellow Must		CQQWLHKQAMQSGSGPS-----PQGPPQRPPLLQCCNELH
tr Q9SQG5	Ara h 6		CEQHIMQRIHQEQEYDSYNFG-----STRSSDQQQRCCDELN
sp P19594	Soybean 2S		CEKHIMEKIQGRGDDDDDDDNHILRTMRGRINYIRRNEGKDEDEEEEHGMQKCCTEMS
sp P23110	SFA-8		CGMYLMKNLGE-----RSQVSPRMREEDHKQLCCMQLK
sp P80356	Avenin		CRQFLVQQCSPVAVPF-----L-----RSQILQQSSCQVMRQQCCRQLE
		*	: ** .:
tr Q9LRC2	Brazil nut		GMD--ESCRCCEGLRMMMRMMQQKEMQP-----
tr EMBL_JRU66866			
	walnut		QMD--EQCQCEGLRQVVRROQQOQGL-----
tn AAK15088	Sesame		NVD-ER CRCEAIRQAVRQQQQEGGY-----
sp P24565	Napin		QED--QVCVCPTLKQAASVVRVQ-----GQH-----
sp P80207	Orient Must		QEE--PLCVCPTLGASKAVKQQIRQQGQQQ--GQQ-----
sp P15322	Yellow Must		QEE--PLCVCPTLGASKAVKQQVRQQLEQQ--GQQ-----
tr Q9SQG5	Ara h 6		EMENTQRCMCALQQIMENQCDGL-----
sp P19594	Soybean 2S		ELR-SPKCQCKALQKIMENQSEEL-----
sp P23110	SFA-8		NLDE--KCMCPAIIIMMLNEPM-----WIR--MRDQVMSMA-----
sp P80356	Avenin		QIPE--QLRCPAIHSVVQAIIMQQQQFFQFPQ--MQQQFFQFPQMQQVT-QGIFQFPQMQQVT
		*	
tr Q9LRC2	Brazil nut		-----RGEQMRRMRLAENIPSRCNLSPMRCPMGGSIAGF-----
tr EMBL_JRU66866			
	walnut		-----RGEEMEEMVQSARDLPNECGISSQRCEIRRSWF-----
tn AAK15088	Sesame		-----QEGQSQQVYQARADLP RR CNMRPQQCQFRVIFV-----
sp P24565	Napin		-----GP--FQSTRIYQIAKNLPNV CNMKQIGTCPFIAIPFFP-----
sp P80207	Orient Must		-----GQQLQHEISRIYQTATHLP RV CNIPRVSCIPFQKTMPGPS-----
sp P15322	Yellow Must		-----GP--HVISR IYQTATHLPKV CNIPQVSVC PFKKTMPGPS-----
tr Q9SQG5	Ara h 6		-----QDRQM VQHFKREL MNLP QQ CNFGAPQRCDLDVSGSRC-----
sp P19594	Soybean 2S		-----EEKQK-KKMEKELINLATMCR FGP MIQC DLSDD-----
sp P23110	SFA-8		-----HNLP IE CNLMSQPCQM-----
sp P80356	AVE3_AVESA		QGIFQFPQLQQVTQ G IF QPQM QGQIE GMRAFALQALPAMCDVYVPPHC PVATAP LG GF-----

A



B

Figure 8 (Continued).

pattern of the four conserved disulphide bonds. However, their protein sequences are only distantly related.

Allergens

2S albumins were among the first plant proteins to be described as allergens over 50 years ago (Youle and Huang, 1981). One of the best characterized is the 2S albumin from yellow mustard seed, Sin a 1 (Menendez-Arias et al., 1988), although others now have been identified in peanut (Burks et al., 1992a), oriental mustard (Monsalve et al., 1993), Brazil nut (Pastorello et al., 1998b), almond (Poltronieri et al., 2002), sunflower (Kelly et al., 2000), and most recently, sesame seed (Pastorello et al., 2001b). It has also been reported that the 2S albumins of soy (Shibasaki et al., 1980) and chickpea (Vioque et al., 1999) are allergens. The IgE epitopes of the allergenic peanut 2S albumin, Ara h 2, have been characterized, three of which are immunodominant (Stanley et al., 1997). Epitope analysis of the 2S allergens from oriental and yellow mustards has identified a major antigenic determinant spanning residues 55–69 of Bra j 1 e (Monsalve et al., 1993), while studies using recombinant Brazil nut 2S albumin have shown the major IgE binding region to be located within the large subunit (Oommen et al., 2000).

Like many other food allergens, the 2S albumins are apparently resistant to proteolysis. Thus, Sin a 1 shows remarkable resistance to proteolysis, with the loss of only a M_r 1,300 fragment after 24 h of trypsinolysis, despite being rich (15.5%) in arginine and lysine residues. These proteins also exhibit an ability to bind or associate with lipids, although not through the action of a binding pocket, as is the case for members of the ns LTP family, but through their activity as efficient emulsifiers (Gueguen et al., 1996). Indeed, the allergenicity of soy 2S albumins appears to be enhanced through interactions with oxidized lipids (Doke et al., 1989). Sin a 1 has also been shown to associate with phospholipid vesicles (Onaderra et al., 1994) perturbing the bilayer and causing the formation of leaky structures. Onaderra et al. (1994) suggested that these properties might play a role in the allergenic properties of mustard 2S albumins by increasing gut permeability, thereby, enhancing allergen uptake and access to the mucosal immune system.

It seems highly likely that the 2S albumin allergens, common with other prolamin superfamily members, are thermostable. However, to date, they have not been characterized in this respect, even though the plant seeds from which they originate are often consumed in a cooked form.

THE CUPIN SUPERFAMILY

The cupins are a superfamily of proteins with a common barrel-like architecture that has variously been described as a “double-stranded β -helix” (Murzin et al., 1995) or a “jellyroll.” However, it was not clear whether these folds shared a common ancestor until the shared sequence motifs, $[G(X)_5HXH(X)_{11}G]$ and $[G(X)_5P(X)_4H(X)_3N]$, (where X is any amino acid residue

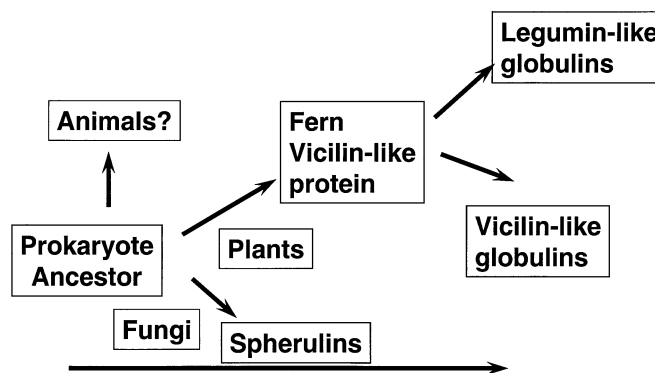


Figure 9 Evolution of cupin superfamily from a common ancestor.

and is a metal binding site in many, but not all, members of the superfamily) were identified (Dunwell and Gane, 1998; Gane et al., 1998; Dunwell, 1998). The term “cupin” (based on the Latin name for a small cask or barrel) was introduced to describe this superfamily, the members of which are probably related by divergent evolution from a common ancestor (Figure 9).

The basic structural feature of the cupins is the presence of a double stranded β -helix domain. Proteins may contain either one or two copies of this, the latter being termed bicupins. The best characterized example of the former is germin, a protein initially identified as being synthesized during the early stages of wheat embryo germination, but now known to exhibit enzymic activity as oxalate oxidase. Each germin monomer contains a single double-stranded β -helix domain fused to a short α -helical domain, and six monomers form a homohexamer that is, in effect, a “trimer of dimers” (Woo et al., 2000).

It is not known whether the duplication of the ancestral germin-like protein has occurred once or several times during the evolution of the cupin superfamily. Hence, the relationships between the various bicupins is obscure. The most widely studied bicupins are the 7S and 11S storage globulins, with the 7S globulin monomer corresponding in structure to the germin dimer and the trimeric 7S globulin to the germin homohexamer (Woo et al., 2000).

While there is one report of a germin being an allergen (Leitner et al., 1998), some notable plant food allergens belong to either the 7S vicilin-like or 11S legumin-like seed storage globulin families. The properties of these allergens are summarized in Table 2; their occurrence, properties, and structures in relation to their allergenic properties are discussed below.

Storage Globulins

Occurrence and Properties

Globulins are generally highly abundant in nuts and seeds, comprising up to 70% of seed protein (Bewley and Black, 1994). They are found in many plant species, including legumes (soy, peas, beans, lentils, chickpea, and peanut) and in nuts/seeds, such as Brazil nut, almonds, sesame, and sunflower. While some

species (such as peanut and soy) synthesize both legumin-like 11S and vicilin-like 7S globulins, others contain only 11S (e.g., Brazil nut) or 7S globulins (e.g., French bean).

In general, these vicilin-like and legumin-like seed globulins exhibit a high degree of thermostability, requiring temperatures in excess of 70°C for denaturation and have a propensity to form large aggregates on heating that still retain, to a large degree, their native secondary structure (Huang et al., 1998; Mills et al., 2001; Yamauchi et al., 1991). At high protein concentrations, these proteins can form heat-set gels (Yamauchi et al., 1991) and may undergo further cross-linking as a consequence of Maillard reactions (Chung and Champagne, 1999).

Structure and Sequence

Three-dimensional structures have been determined for two 7S globulins: phaseolin from French bean (Lawrence et al., 1990, 1994) and canavalin from jack bean (Ko et al., 1993, 2000). Figure 10 shows the canavalin trimer (Ko et al., 2000), with the basic cupin fold of domain 1 comprising a number of β -strands and α -helices forming domain 2. Single cupin folds are found in other proteins, such as oxalate oxidase, gene duplication, and fusion, giving rise to the 7S bicupin with an approximate 2-fold axis (Figure 10). Canavalin and phaseolin are structurally very similar proteins, especially if the *N*- and *C*-terminal halves are

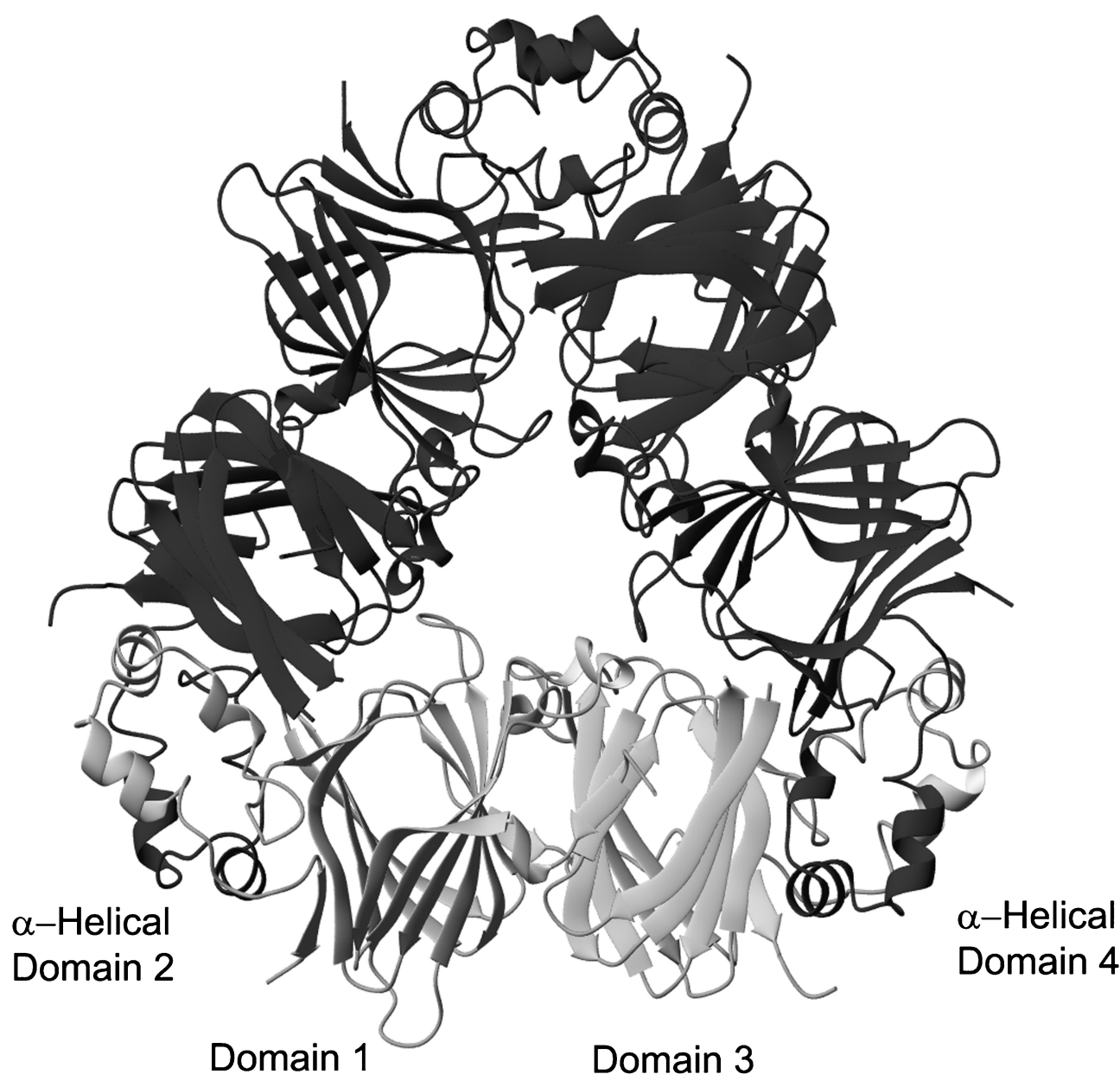


Figure 10 Structure of canavalin, a 7S globulin belonging to the cupin superfamily, from seeds of jack bean. Ribbon representation based on the structure of Ko et al. (1993, 2000).

superimposed separately, with r.m.s. distances of 0.49–0.52 Å for the 112–125 C α pairs in the *N*-terminal domain and 0.49–0.51 Å for 115–131 pairs in the *C*-terminal domain (Ko et al., 1993, 2000). The crystal structure of the trimeric proglycinin a(1a)B(1b) (11S globulin precursor) protein of soy (Adachi et al., 2001) has revealed a structure similar to that of the 7S globulins. However, whereas the two 7S globulin structures contain some disordered regions, these represent approximately 20% of the structure in the 11S proglycinin. Adachi et al. (2001) suggested that these disordered regions prevent formation of the native glycinin hexamer until proglycinin is proteolytically processed to glycinin.

Allergens

The 11S and 7S globulins of soy were among the first plant food allergens to be characterised (Shibasaki et al., 1980), although interpretation of the results of these early investigations is complicated by the cross-contamination of globulin fractions. More recent studies have unequivocally confirmed the allergenic activity of the 7S globulin of soy (β -conglycinin) (Burks et al., 1988) and subsequently the 7S globulins of peanut (conarachin, also known as Ara h 1) (Burks et al., 1991), walnut (Jug r 2) (Teuber et al., 1999), sesame (Beyer et al., 2002), cashew nut (Ana c 1) (Wang et al., 2002), and one of the subunits of the proteolytically processed 7S globulin of lentil (Sanchez-Monge et al., 2000). The 11S globulins have also been confirmed as allergens in peanut, (arachin, Ara h 3) (Burks et al., 1992b), soy (glycinin) (Rabjohn et al., 1999) and almond (also known as almond major protein, AMP) (Roux et al., 2001). They have also been implicated as allergens in coconut and walnut (Teuber and Petersen, 1999).

Detailed epitope mapping studies of Ara h 1 have shown the IgE-epitopes to be located in the α -helical regions of the protein that contain a high proportion of aromatic residues and form the hooks that hold the subunits into the trimeric structure of conarachin (Shin et al., 1998). IgE epitopes in the 11S globulin allergens of peanut (arachin, Ara h 3) and soy (glycinin) are located in the equivalent domain on the acidic subunits of these proteins (Rabjohn et al., 1999; Beardslee et al., 2000), although only one epitope is common to Ara h 3 and soya 7S globulin (Helm et al., 2000a, b; Beardslee et al., 2000). Some 7S globulins are glycosylated and, for Ara h 1 at least, these glycosylated sites are significant IgE epitopes and may represent another source of thermostable epitopes (van Ree et al., 2000). However, the role of glycosylated epitopes in eliciting allergic reactions *in vivo* is still a matter of debate (Alberse, 1998).

The globulins also show resistance to proteolysis, Ara h 1, retaining both its trimeric structure and IgE reactivity following pepsin digestion (Maleki et al., 2000). Indeed, the 11S globulin of soy (glycinin) has been shown to form stable intermediates of Mr $\sim$ 280,000, known as glycinin T on trypsinolysis. These intermediates retain the quaternary structure of glycinin and result from proteolytic clipping of the acidic subunits to form Mr 13,000 and 16,000 fragments, the basic subunits re-

maining intact. Similar intermediates are formed by following chymotrypsinolysis (Kamata and Shibasaki, 1978; Plumb et al., 1994). Resistance to denaturation and proteolysis is a general characteristic of members of the cupin superfamily. For example, the isolation procedure for oxalate oxidases from wheat bran involves digestion with pepsin and boiling for 30 m (Lane, 2000). It seems likely that such exceptional stability may play a role in determining the allergenic activity of globulin storage proteins.

CYSTEINE PROTEASES

Cysteine proteases of the C1, or papain-like, family are part of a much larger family of cysteine proteases that were originally characterized by their having a cysteine residue as part of their catalytic site. The C1 family was identified as having conserved Gln, Cys, His, and Asn residues at the active site and includes many endopeptidases, aminopeptidases, and dipeptidyl peptidases, some enzymes having both both exo- and endopeptidase activities (Rawlings and Barrett, 1993). While some family members share sequence homology, they may not have protease activity, a notable example being the P34 protein from soybean, where the active site cysteine residue has been replaced with a glycine (Kalinski et al., 1990). Members of the papain family are widespread in nature, having been found in baculovirus, eubacteria, yeasts, and practically all protozoa, plants, and mammals. The proteins are generally lysosomal or secreted, being synthesized as a proprotein, proteolytic cleavage of the propeptide being required for enzyme activation. Many members are also glycosylated, with the *N*-linked glycan moiety of the cysteine protease of pineapple, bromelain, having been determined to be Man α -D-Man (1–6)-[β -D-Xyl-(1–2)]- β -D-Man-(1–4)- β -D-GlcNAc-(1–4)-[α -L-Fuc-(1–3)] β -D-GlcNAc- (1–N) (Ishihara et al., 1979).

One important role for cysteine proteases in eukaryotic systems that has emerged in recent years is their critical role in the programmed cell death (apoptosis) pathway of animal cells (Martin and Green, 1995). Such a role is now emerging in plants (Solomon et al., 1999) and is associated with apoptotic events that accompany plant responses to pathogen attack and environmental stress.

A number of allergens have been identified belonging to the C1 protease family, including inhalant allergens such as the dust mite allergen Der p 1, a β -expansin with a papain-like domain which is a major allergen of timothy grass pollen, Phl p 1 (Grobe et al., 1999) and a number of food allergens, notably those from kiwi fruit and soybean. There is evidence that the protease activity of the dust mite allergen, Der p 1, is important in its allergenicity, as it can cleave the human IgE receptor, CD23, and may, thereby, ablate the feedback mechanism that normally regulates IgE synthesis (Hewitt et al., 1995; Schulz et al., 1995). Data from animal models indicate that this may also be true of papain (Chambers et al., 1998).

The properties of the food allergens belonging to this superfamily are summarized in Table 2, while their occurrence,

properties, and structures in relation to their allergenic properties, are discussed below.

Occurrence and Properties

Cysteine proteinases occur widely in plants where they may participate in a range of processes, such as senescence of flowers and leaves (Becker et al., 1994; Drake et al., 1996; Jones et al., 1995), processing of enzyme precursors (Gietl et al., 1997), and digestion of storage proteins during seed germination (de Barros and Larkins, 1990; Fischer et al., 2000; Mikkonen et al., 1996). The cysteine proteinases in maize have also been associated with resistance to lepidopteran pests (Jiang et al., 1995; Pechan et al., 2000), while the high levels of cysteine proteinases in fruits (papaya, fig, pineapple, and kiwi) are also consistent with a protective role.

The best characterized cysteine proteinase of plant origin is papain, which is present in the latex of papaya (*Carica papaya*). Preparations of papain are widely used in the food industry, including to supplement malt enzymes in brewing and to tenderize meat. They are also used therapeutically, as an anti-inflammatory drug. Related proteinases are present in kiwi fruit (actinidin), fig latex (ficin), pineapple stem (bromelain), and soybean seeds (34 kD oil body-associated protein, P34).

Structure and Sequence

Sequence comparisons show a high degree of relatedness across the family, regarding the residues surrounding the catalytic site. Thus, in papain, the catalytic residues are Cys-25 and His-159. Other important residues are Gln-19, which helps form the 'oxyanion hole,' and Asn-175, which orientates the imidazole ring of His-159. Despite these structural similarities, the C proteases have only low levels of overall sequence homology; the dust mite allergen, Der p 1 having only approximately 30% identical to the plant cysteine proteases (Figure 11).

The mature enzymes are generally $M_r \sim 25,000$ proteins. Although the gene sequences are often considerably longer, the signal sequence is followed by an α -helical prodomain, which is often necessary for folding. It also functions to inactivate the proenzyme by blocking the active site. Proteolytic processing is required to activate the enzyme (Groves et al., 1996; Coulombe et al. 1996). Three dimensional structures have been determined for a number of C1 proteases, including actinidin (Baker, 1980; Figure 12) and papain (Kamphuis et al., 1984). Both have two domain structures, with an α -helical N-terminal domain and an $\alpha + \beta$ C-terminal domain that both contribute residues to an active site cleft that lies between the domains.

Allergens

The best characterized cysteine protease food allergen is actinidin (Act c 1), the major allergen from kiwi fruit (Fahlbusch et al., 1998; Möller et al., 1997; Pastorello et al. 1998a). Unlike other kiwi allergens, actinidin shows no IgE cross-reactivity with

pollen proteins, implying that it acts to sensitize individuals via the GI tract, rather than through inhalation (Gall et al., 1994; Pastorello et al., 1996). Cysteine proteases have also been identified as allergens in pineapple (the enzyme bromelain) (Bauer, 1979) and papaya (papain); the protease of figs, ficin, having been implicated as an allergen (Diez-Gomez et al., 1998). Patients with kiwi allergy also show IgE reactivity to papain and bromelain (Gall et al., 1994; Pastorello et al., 1998a), while hypersensitivity to the sap or latex from the weeping fig, *Ficus benjamina*, results in cross-reactivity with fig fruit and with papain (Diez-Gomez et al., 1998). Allergies to bromelain and papain may be triggered by dermal sensitization, allergies to these proteins are associated with their use in cosmetics, or orally through medical applications (Zentner et al., 1997; Soto-Mera et al., 2000). Papain may also sensitize by inhalation, as it has also been reported as a pollen allergen (Blanco et al., 1998) and also through occupational exposure (Freye, 1988). One of the major IgE-binding epitopes on bromelain is an N-linked glycan, which may be responsible for the cross-reactive nature of this allergen (Afferni et al., 1999).

Another food allergen belonging to the cysteine protease superfamily is the soybean 34 kD, so-called oil body-associated protein, known variously as Gly m Bd 30 K, Gly m 1, or P34. The major allergen involved in soybean-induced atopic dermatitis (Ogawa et al., 1993) is actually associated with the storage vacuoles of soybean (Kalinski et al., 1992). Traces of the 46 kDa precursor are found in the seed, with an M_r 32,000 species formed, following soybean germination. Gly m Bd 30 K exists as complex of molecular weight $>300,000$, as it becomes disulphide-linked to the α , α' subunits of the soya globulin β -conglycinin (Ogawa et al., 1993; Samoto et al., 1996) with at least one of the two putative N-glycosylation sites being glycosylated (Bando et al., 1996). Serum reactive linear epitopes have been reported from Gly m Bd 30 K (Helm et al., 1998) and the five immunodominant epitopes were analysed further by mutation, some of the epitopes reported corresponding to the pro-region of the protein which is present only in the precursor proteins which is found only at low levels in mature seed (Helm et al., 2000).

There is little information available regarding the stability characteristics of any of these allergens, although, as they are typically stabilized by three disulphide bridges, they are moderately thermostable, as has been demonstrated for papain which requires heating to 60°C before becoming denatured (Sumner et al., 1993). While the levels of actinidin have been shown to vary widely between cultivars of kiwi fruit, suggesting that breeding may be a way of reducing the allergens in these fruits (Nishiyama and Oota, 2002), there appears to be little cultivar variation in Gly m Bd 30 K in soybeans (Yaklich et al., 1999).

OTHER ALLERGENS

While the majority of food allergens thought to sensitize via the GI tract belong to the prolamin, cupin, and C1

sp	P00785	Actinidin	-----MGLPKSFVMSLLFFSTLLILSLAFNAKNTLQRTNDEVKAMYE
sp	P00784	Papain	MAMIPSIKLLFVAICLFVYMGLSFGDF-----SIVGYSQNDLTSTERLIQLFE
tr	O23791	Bromelain	MASKVQLVFLFLFLCAMWASPSAASRD-----EPNDPMMKRFE
sp	P22895	Soybean P34	-----MGFLVLLLFSLGLSSSSSISTHRSILDLDLTKEFTTQKQVSSLFQ
: : :			
sp	P00785	Actinidin	SWLIKYGKSYNSLGEWERRFEIFKETLRFIDEHN---ADTNRSYKVGILNQFADL
sp	P00784	Papain	SWMLKHNKIYKNIDEKIYRFEIFKDNLKYIDETN---KKN-NSYWLGLNVFADM
tr	O23791	Bromelain	EWMAEYGRVYKDDDEKMRRFQIFKNNVKHIETFN---SRNENSYTLGINQFTDM
sp	P22895	Soybean P34	LWKSEHGRVYHNHEEEAKRLEIFKNNSNYIRDMNANRKSP-HSHRLGLNKFADI
* :.: *.. * *::*:.. . * * .: :*: *:::			
sp	P00785	Actinidin	TDEEFIRSTYLGFTSGSNKTKV--SNRYEPR--FGQVLPSYVDWRSAGAVVDIK
sp	P00784	Papain	SNDEFKEKYTGSIAGNYTTTE-LSYEEVLND--GDVNIPEYVDWRQKGAVTPVK
tr	O23791	Bromelain	TKSEFVAQYTGVSPLNIERE--PVVSFDDV--NISAVPQSIDWRDYGAVNEVK
sp	P22895	Soybean P34	TPQEFSSKKYLQAPKDVSQQIK-MANKMKKKEQYSCDHPPASWDWRKKGVITQVK
: .** * * *** *: :			
sp	P00785	Actinidin	SQGECCGCWAFSAIATVEGINKIVTGVLIISLSEQELIDCGRTQNTRGCGNGGYIT
sp	P00784	Papain	NQGSCGSCWAFSAVVTIEGIIKIRTGNLNEYSEQELLDCCR--RSYGCNGGYPW
tr	O23791	Bromelain	NQNPCGSCWSFAAIATVEGIYKIKTGVLVSLSEQEVLDCAV---SYGCKGGWVN
sp	P22895	Soybean P34	YQGGCGRGWAFSATGAIEAAHAIAITGDLVSLSEQELVDCVE--ESEGSYNGWQY
* . * *::: :*. * * * . *****: ** : * . *:			
sp	P00785	Actinidin	DGFQFIINNGGINTEENYPYTAQDGECLDLQNEKYVTIDTYENV-----PY
sp	P00784	Papain	SALQLV-AQYGIHYRNTYPYEGVQRYCRSREKGPYAAKTGQVVRQV-----QP
tr	O23791	Bromelain	KAYDFIISNNGVTTEENYPYLAYQGTNCANSFPNS-AYITGYSYV-----RR
sp	P22895	Soybean P34	QSFQVLEHGGIATDDDDYPYRAKEGRCKANKIQD-KVTIDGYETLIMSDESTES
. : : : * : *** . : * . . :			
sp	P00785	Actinidin	NNEWALQTAVTYQPVSVVALDAAGDAFKHYSSGIFTGPCGTA---IDHAVTIVGY
sp	P00784	Papain	YNEGALLYSIANQPVSVVLEAAGKDFQLYRGGIFVGPCGNK---VDHAVAAGVY
tr	O23791	Bromelain	NDERSMMYAVSNQPIAALIDAS-ENFQYYNGGVFSGPCGTS---LNHAITIIGY
sp	P22895	Soybean P34	ETEQAFLSAILEQPISVSIDA--KDFHLYTGGIYDGENCTSPYGINHFVLLVGY
* : : : **::: : : * * .::: * . :*: : **			
sp	P00785	Actinidin	GT-EGGIDYWIVKNSWDTTWGEEGYMRILRNVGGA-GTCGIATMPSYPVKYNNQ
sp	P00784	Papain	GP-----NYILIKNSWGTGWGENGYIRIKRGTGNSYGVCGLYTSSFYPVKN---
tr	O23791	Bromelain	GQDSSGTYKWIVRNSWGSSWGEGGYVRMARGVSSSSGVCGIAMAPLFTLQSGA
sp	P22895	Soybean P34	GS-ADGVDYWIAKNSWGFWDGEDGYIWIQRNTGNLLGVCGMNYFASYPTKEESE
* * : :***. *** *: : * .. *.** : . :*			
sp	P00785	Actinidin	NHPKPYSSLINPPAFSMSKDGVPVGDGQRYSA
sp	P00784	Papain	-----
tr	O23791	Bromelain	NAEVIKMVSET-----
sp	P22895	Soybean P34	TLVSARVKGHRRVDHSPL-----

Figure 11 Aligned sequences of cysteine protease C1 family members from plants compared with that of the dust mite allergen Der p 1. Allergens are shown in bold. * conserved residues; . conserved substitutions; : semi-conserved substitutions. Conserved residues around the catalytic site are shaded. Allergen names are preceded by database accessions: sp—Swiss Prot; tr—TREMBL.

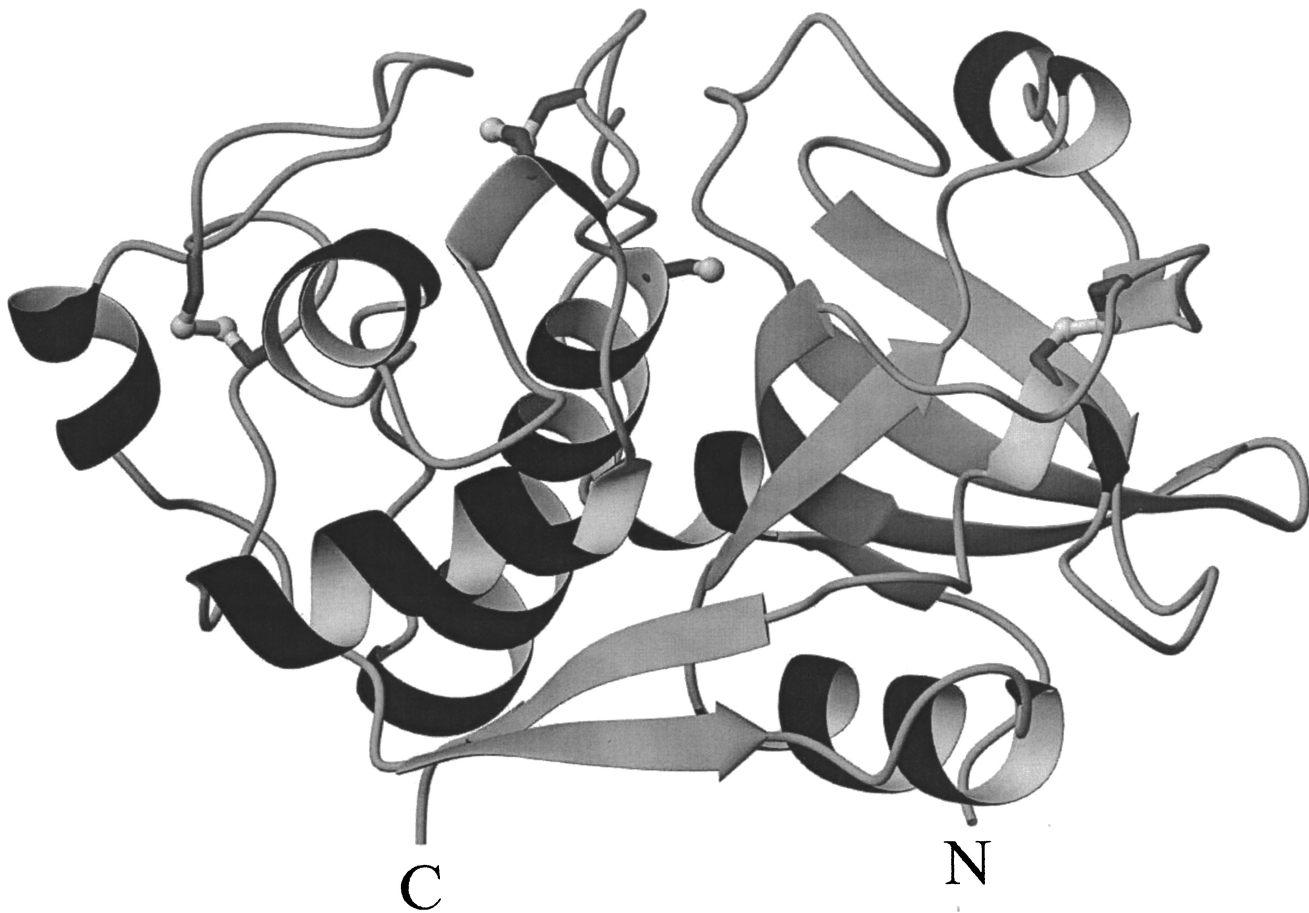


Figure 12 Structure of the cysteine protease, actinidin, a major allergen of kiwi fruit. Ribbon drawing of actinidin (Baker, 1980 and PDB code 2ACT) with helices dark on one side of the ribbon, while beta-sheets are represented with light grey arrows. The view is into the protease active site, with the catalytic cysteine near the centre. The catalytic cysteine's side chain and the three disulphides are shown as ball and stick drawings. The *N*-terminal helical domain is to the left and the C-terminal $\alpha + \beta$ domain to the right. The *N*- and C-termini are labelled.

protease superfamilies, the remaining food allergens (many considered minor allergens) are scattered across a wide range of protein families. This includes the Kunz inhibitor family, a group of protease inhibitors found in plants that can inhibit serine-, thiol-, and aspartyl- proteases and probably play a role in plant defence. They have been identified as allergens in soybean (Moroz and Yang, 1980; Herian et al., 1980; Burks et al., 1994) and potato (Seppala et al., 2001). It is thought that their stability to processing and digestion is important for their allergenic activity. The peanut lectin, agglutinin, has also been identified as a minor allergen in peanut (Burks et al., 1994). A number of other allergens may sensitize via the GI tract, although, at present, there may be a lack of clinical data to support this supposition. These include the thaumatin-like protein allergens of fruits, such as cherry and apple, which are small, highly disulphide-bonded proteins involved in the protection of plants against pathogens, for which no pollen analogues have been identified. They may be involved in oral sensitization to fruits and vegetables (Hoffmann-Sommergruber, 2000).

DOES PROTEIN STRUCTURE AND PROPERTIES PLAY A ROLE IN PLANT FOOD ALLERGENICITY?

The apple allergen Mal d 1, involved in the pollen-fruit allergy syndrome, is considered to be an "incomplete" allergen, i.e., ingestion of apple fruit does not stimulate the formation of anti-Mal d 1 IgE, but does trigger histamine release from mast cells loaded with IgE to the birch pollen allergen Bet v 1 to which Mal d 1 is homologous (Alberse, 2000). Such a thermolabile, readily digested protein is, therefore, able to elicit an allergic reaction, but is unable to sensitize an individual through ingestion. Implicit in this observation is that for a food allergen to sensitize an individual via the gastrointestinal tract, it must possess certain structural and biological attributes that preserve its structure from the destructive effects of the gastrointestinal tract, including resistance to low pH, proteolysis, and surfactants, such as bile salts. Two properties that seem, in general, to be shared by plant protein allergens are abundance coupled with structural stability. This combination of properties may help to ensure that enough of the protein survives in a sufficiently intact form to be

taken up by the gut and sensitize the mucosal immune system. Similar properties are shared by food allergens of animal origin, with the milk allergens, such as cow's milk β -lactoglobulin (Wal, 1998), cod parvalbumin (Bugajska-Scretter et al., 1998), and hen egg ovomucoid (Cooke and Sampson, 1997) all being abundant, thermostable proteins that are resistant to proteolysis. An exception to this is the ns LTP family, which is not abundant in fruits *per se*, but that do make up a significant proportion of the protein in some plant tissues, such as fruit peel.

Several factors are thought to contribute to protein stability, including a compact 3-D structure, ligand binding and glycosylation. Protein engineering approaches to improve protein thermostability indicate that while stability is generally associated with increased numbers of disulphide bridges, the relationship is complex. Thus, disulphide bonds can contribute to the stability of both the native and denatured form of a protein, the overall protein stability arising from a balance between the stabilities of these different states (Zavodsky et al., 2001). Protein glycosylation also plays a role, and glycosylation. For example, the 7S globulin of peas increases trimer stability and resistance to chemical denaturation (Pedrosa et al., 2000). The relationship between protein stability and susceptibility to proteolysis is less clear, although proteolytic cleavage sites are clearly more accessible to proteases when presented in more mobile regions, which are generally found on the surface of a protein. About 20% of the proglycinin structure is mobile (Adachi et al., 2001), and it may be that the globulins are susceptible to proteolytic cleavage, but are able to retain much of their quaternary structure following digestion.

Another property shared by many plant food allergens is an ability to bind lipids and associate with membranes. This property is also shared by allergens of animal origin, and of non-food allergens, such as those found in parasites (Moore et al., 1999). Thus, several members of the lipocalin superfamily, a group of proteins able to bind small lipophilic molecules, are allergens, including β -lactoglobulin from cow's milk, inhalant allergens such as mouse urinary protein, and lipocalins found in animal dander (Mantylarvi et al., 1991). Proteins, such as the lipocalins and ns LTPs, which possess a lipid-binding pocket, show increased stability when the pocket is occupied. Thus, the thermostability of β -lactoglobulin has been found to increase on lipid binding (Creamer, 1995). Many other plant food proteins belonging to the prolamin and cupin superfamilies will also associate with lipid structures, either binding to membranes or acting as emulsifiers, by adsorbing to lipid droplet surfaces. Such associations may occur either in the food as a consequence of processing, or in the gastrointestinal tract as a consequence of the digestive process. Such lipid-associated proteins may be protected from the degradative environment in the gut, and there is evidence that proteins associated with lipid bilayers have enhanced thermal stability (Taneva et al., 2000). It is also possible that lipid-binding may assist in allergens being taken up by the gut mucosa, either through association with cell membranes, as has been suggested for 2S albumins, or by co-absorption on the surface of lipid droplets.

Thus, the structural features and properties of proteins belonging to the prolamin and cupin superfamilies may play a role in predisposing these proteins to becoming allergens. However, it is clear that these properties are not the sole determinant of allergenic potential. As yet, it remains to be shown why in one plant food, such as Brazil nut, the major allergen is the 2S albumin, and not the 11S globulin, while in soy, the converse appears to be the case. The immune responses generated in both allergic disease and following parasitic infection are similar, with Th2-mediated responses dominating, the consequent production of IgE and eosinophilia being characteristic of both conditions. There is evidence that atopy is promoted by genetic variants within the Th2 immune system, and there are also indications that one haplotype of the major histocompatibility complex is associated with immune responsiveness to the nematode allergen NBA in mice, rats, and man. This has led to the suggestion that there is a counterbalancing evolutionary benefit, with genetic variants conferring an increased risk of atopy also having conferred increased biological fitness through protection against helminth infection (Mao et al., 2000). Therefore, the realization of the allergenic potential of a protein also depends on genetic factors that determine the immune responsiveness of an individual may be paramount.

ACKNOWLEDGEMENT

IACR and IFR receive grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom. Part of this work was supported by EU FAIR CT-98-4356.

REFERENCES

- Adachi, M., Takenaka, Y., Gidamis, A.B., Mikami, B., and Utsumi, S. 2001. Crystal structure of soybean proglycinin A1aB1b homotrimer. *J. Mol. Biol.*, **12**:291–305.
- Adachi, T., Izumi, H., Yamada, T., Tanaka, K., Takeuchi, S., Nakamura, R., and Matsuda, T. 1993. Gene structure and expression of rice seed allergenic proteins belonging to the α -amylase/trypsin inhibitor family. *Plant Mol. Biol.*, **21**:239–248.
- Afferri, C., Iacovacci, P., Barletta, B., Di Felice, G., Tinghino, R., Mari, A., and Pini, C. 1999. Role of carbohydrate moieties in IgE binding to allergenic components of *Cupressus arizonica* pollen extract. *Clin. Exp. Allergy*, **29**:1087–1094.
- Alberse, R.C. 1998. Clinical relevance of carbohydrate allergen epitopes. *Allergy*, **53**(45 Suppl):54–57.
- Alberse, R.C. 2000. Structural biology of allergens. *J. Allergy Clin. Immunol.*, **106**:228–238.
- Alvarez, A.M., Adachi, T., Nakase, M., Aoki, N., Nakamura, R., and Matsuda, T. 1995. Classification of rice allergenic protein cDNAs belonging to the α -amylase/trypsin inhibitor gene family. *Biochim. et Biophys. Acta*, **1251**:201–204.
- Anderson, J.A. 1996. Allergic reactions to foods. *Crit. Rev. Food Sci. Nutr.*, **36**(Suppl):S19–38.
- Anderson, O.D., Greene, F.C., Yip, R.E., Halford, N.G., Shewry, P.R., and Malpica-Romero, J.-M. 1989. Nucleotide sequences of the two high-molecular-weight glutenin genes from the D-genome of a hexaploid bread

- wheat, *Triticum aestivum* L. cv Cheyenne. *Nucleic Acids Research*, **17**:461–462.
- “The Arabidopsis genome initiative” 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**:796–815.
- Asero, R., Mistrello, G., Roncarolo, D., de Vries, S.C., Gautier, M.F., Ciurana, C.L., Verbeek, E., Mohammadi, T., Knul-Brettlova, V., Akkerdaas, J.H., Bulder, I., Aalberse, R.C., and van Ree, R. 2000. Lipid transfer protein: a pan-allergen in plant-derived foods that is highly resistant to pepsin digestion. *Int. Arch. Allergy Immunol.*, **122**:20–32.
- Asero, R., Mistrello, G., Roncarolo, D., Amato, S., and van Ree, R. 2001. A case of allergy to beer showing cross-reactivity between lipid transfer proteins. *Ann. Allergy Asthma Immunol.*, **87**:65–67.
- Baker, E.N. 1980. Structure of actinidin after refinement at 1.7 Å resolution. *J. Mol. Biol.*, **141**:441–484.
- Bando, N., Tsuji, H., Yamanishi, R., Nio, N., and Ogawa, T. 1996. Identification of the glycosylation site of a major soybean allergen, Gly m Bd 30K. *Biosci. Biotechnol. Biochem.*, **60**:347–8.
- Bartels, D., Altosaar, I., Harberd, N.P., Barker, R.F., and Thompson, R.D. 1986. Molecular analysis of γ -gliadin gene families at the complex *Gli-1* locus of bread wheat (*T. aestivum* L.). *Theor. Appl. Genet.*, **72**:845–853.
- Baur, X. 1979. Studies on the specificity of human IgE-antibodies to the plant proteases papain and bromelain. *Clin. Allergy*, **9**:451–457.
- Bauer, L., Ebner, C., Hirschwehr, R., Wuthrich, B., Pichler, C., Fritsch, R., Scheiner, O., and Kraft, D. 1996. IgE cross-reactivity between birch pollen, mugwort pollen and celery is due to at least three distinct cross-reacting allergens: Immunoblot investigation of the birch-mugwort-celery syndrome. *Clin. Exp. Allergy*, **26**:1161–1170.
- Beardslee, T.A., Zeece, M.G., Sarath, G., and Markwell, J.P. 2000. Soybean glycinin G1 acidic chain shares IgE epitopes with peanut allergen Ara h 3. *Int. Arch. Allergy Immunol.*, **123**:299–307.
- Beccari De Frumento. 1745. *De Bononiensi Scientiarum et Artium Instituto atque Academia Commentarii*, II., Part I, pp. 122–127.
- Bech, L. M., Vaag, P., Heinemann, B., and Breddam, K. 1995. Throughout the brewing process barley lipid transfer protein 1 is transformed into a more foam-promoting form. In: *Procc. Congr. Eur. Brew. Conv. Lisbon*, pp. 561–568.
- Becker, C., Fischer, J., Nong, V.H., and Müntz, K. 1994. PCR cloning and expression analysis of cDNAs encoding cysteine proteinases from germinating seeds of *Vicia sativa* L. *Plant Mol. Biol.*, **26**:1207–1212.
- Behnke, C.A., Yee, V.C., Le Trong, I., Pedersen, L.C., Stenkamp, R.E., Kim, S.S., Reeck, G.R., and Teller, D.C. 1988. Structural determinants of the bifunctional corn Hageman factor inhibitor: X-ray crystal structure at 1.95 angstrom resolution. *Biochem.*, **37**:15277–15288.
- Bewley, J.D., and Black, M. 1994. *Seeds, Physiology of Development and Germination*, 2nd Edition. Plenum Press, New York.
- Beyer, K., Bardina, L., Grishina, G., and Sampson, H.A. (2002). Identification of sesame seed allergens by 2-dimensional proteomics and Edman sequencing: Seed storage proteins as common food allergens. *J. Allergy Clin. Immunol.*, **110**:154–159.
- Blanco, C., Ortega, N., Castillo, R., Alvarez, M., Dumpierrez, A.G., and Carrillo, T. 1998. *Carica papaya* pollen allergy. *Ann. Allergy Asthma Immunol.*, **81**:171–175.
- Boulter, D. and Croy, R.R.D. 1997. The structure and biosynthesis of legume seed storage proteins: A biological solution to the storage of nitrogen in seeds. *Adv. Bot. Res.*, **27**:1–84.
- Breiteneder, H. and Ebner, C. 2000. Molecular and biochemical classification of plant-derived food allergens. *J. Allergy Clin. Immunol.*, **106**:27–36.
- Bush, R.K. and Hefle, S.L. 1996. Food allergens. *Crit. Rev. Food Sci. Nutr.*, **36**(Suppl):S119–163.
- Brenna, O., Pompei, C., Ortolani, C., Pravettoni, V., Farioli, L., and Pastorello, E.A. 2000. Technological processes to decrease the allergenicity of peach juice and nectar. *J. Agric. Food Chem.*, **48**:493–497.
- Bugajska-Scretter, A., Elfman, L., Fuchs, T., Kapiotis, S., Rumpold, H., Valenta, R., and Spitzauer, S. 1998. Parvalbumin, a cross-reactive fish allergen, contains IgE-binding epitopes sensitive to periodate treatment and Ca^{2+} depletion. *J. Allergy Clin. Immunol.*, **101**:67–74.
- Burks, A.W., Cockrell, G., Connaughton, C., Guin, J., Allen, W., and Helm, R.M. 1994. Identification of peanut agglutinin and soybean trypsin inhibitor as minor legume allergens. *Int. Arch. Allergy Immunol.*, **105**:143–149.
- Burks, A.W., Brooks, J.R., and Sampson, H.A. 1988. Allergenicity of major component proteins of soybean determined by enzyme-linked immunosorbent assay (ELISA) and immunoblotting in children with atopic dermatitis and positive soy challenge. *J. Allergy Clin. Immunol.*, **81**:1135–1142.
- Burks, A.W., Williams, L.W., Connaughton, C., Cockrell, G., O'Brien, T.J., and Helm, R.M. 1992b. Identification and characterization of a second major peanut allergen, Ara h III, with use of the sera of patients with atopic dermatitis and positive peanut challenge. *J. Allergy Clin. Immunol.*, **90**:962–969.
- Burks, A.W., Williams, L.W., Helm, R.M., Connaughton, C., Cockrell, G., and O'Brien, T. 1991. Identification of a major peanut allergen, Ara h I, in patients with atopic dermatitis and positive peanut challenges. *J. Allergy Clin. Immunol.*, **88**:172–179.
- Burks, A.W., Williams, L.W., Connaughton, C., Cockrell, G., O'Brien, T.J., and Helm, R.M. 1992a. Identification and characterization of a second major peanut allergen, Ara h II, with use of the sera of patients with atopic dermatitis and positive peanut challenge. *J. Allergy Clin. Immunol.*, **90**:962–969.
- Carbonero, P. and García-Olmedo, F. 1999. A multigene family of trypsin/ α -amylase inhibitors from cereals. In: *Seed Proteins*. pp. 617–633. Eds., Shewry, P.R., and Casey, R., Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Chambers, L., Brown, A., Pritchard, D.I., Sreedharan, S., Brocklehurst, K., Kalsheker, N.A. 1998. Enzymatically active papain preferentially induces an allergic response in mice. *Biochem. Biophys. Res. Commun.*, **253**:837–840.
- Charvolin, D., Douliez, J.-P., Marion, D., Cohen-Addad, C., and Pebay-Peyroula, E. 1999. The crystal structure of a wheat nonspecific lipid transfer protein (ns-LTP1) complexed with two molecules of phospholipid at 2.1 angstrom resolution. *Eur. J. Biochem.*, **264**:562–568.
- Christensen, A.B., Cho, B.H., Naesby, M., Gregersen, P.L., Brandt, J., Madriz-Ordeñana, K., Collinge, D.B., and Thordal-Christensen, H. 2002. The molecular characterization of two barley proteins establishes the novel PR-17 family of pathogenesis-related proteins. *Mol. Plant Pathol.*, **3**:135–144.
- Chung, S.-Y. and Champagne, E.T. 1999. Allergenicity of maillard reaction products from peanut proteins. *J. Agric. Food Chem.*, **47**:5227–5231.
- Colot, V., Robert, L.S., Kavanagh, T.A., Bevan, M.W., and Thompson, R.D. 1987. Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. *EMBO J.*, **6**:3559–3564.
- Cooke, S.K. and Sampson, H.A. 1997. Allergenic properties of ovomucoid in man. *J. Immunol.*, **159**:2026–2032.
- Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort, J.S., and Cygler, M. 1996. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO J.*, **15**:5492–5503.
- Creamer, L.K. 1995. Effect of sodium dodecyl sulfate and palmitic acid on the equilibrium unfolding of bovine β -lactoglobulin. *Biochem.*, **34**:7170–7176.
- Crouch, M.T., Tenbarger, K.M., Simon, A.E., and Ferl, R. 1983. cDNA clones for *Brassica napus* seed storage proteins: Evidence from nucleotide sequence analysis that both subunits of napin are cleaved from a precursor polypeptide. *J. Mol. Appl. Genet.*, **2**:273–283.
- Curioni, A., Santucci, B., Cristado, A., Canistraci, C., Peittravalle, M., Simonato, B., and Giannattasio, M. 1999. Urticaria from beer: An immediate hypersensitivity reaction due to a 10 kDa protein derived from beer. *Clin. Exp. Allergy*, **29**:407–413.
- de Barros, E.G. and Larkins, B.A. 1990. Purification and characterization of zein-degrading proteases from endosperm of germinating maize seeds. *Plant Physiol.*, **94**:297–303.
- Derbyshire, E., Wright, D.J., and Boulter, D. 1976. Legumin and vicilin, storage proteins of legume seeds. *Phytochem.*, **15**:3–24.
- Diez-Gomez, M.L., Quirce, S., Aragonese, E., and Cuevas, M. 1998. Asthma caused by *Ficus benjamina* latex: Evidence of cross-reactivity with fig fruit and papain. *Ann. Allergy Asthma Immunol.*, **80**:24–30.
- Doke, S., Nakamura, R., and Torii, S. 1989. Allergenicity of food proteins interacted with oxidized lipids in soybean-sensitive individuals. *Agric. Biol. Chem.*, **53**:1231–1235.

- Douliez, J.-P., Michon, T., Elmorjani, K., and Marion, D. 2000. Structure, biological and technological function of lipid transfer proteins and indolines, the major lipid binding proteins from cereal kernels. *J. Cereal. Sci.*, **32**:1–20.
- Drake, R., John, I., Farrell, A., Cooper, W., Schuch, W., and Grierson, D. 1996. Isolation and analysis of cDNAs encoding tomato cysteine proteases expressed during leaf senescence. *Plant Mol. Biol.*, **30**:755–767.
- Dunwell, J.M. 1998. Cupins: A new superfamily of functionally diverse proteins that include germins and plant storage proteins. *Biotechnol. Genet. Eng. Rev.*, **15**:1–32.
- Dunwell, J.M. and Gane, P.J. 1998. Microbial relatives of seed storage proteins: Conservation of motifs in a functionally diverse superfamily of enzymes. *J. Mol. Evol.*, **46**:147–154.
- Duro, G., Colombo, P., Costa, M.A., Izzo, V., Porcasi, R., Di Fiore, R., Locorotondo, G., Mirisola, M.G., Cocchiara, R., and Geraci, D. 1996. cDNA cloning, sequence analysis and allergological characterization of Par j 2.0101, a new major allergen of the *Parietaria judaica* pollen. *FEBS Letts.*, **399**:295–295.
- Egorov, T.A., Odintsova, T.I., Musolyamov, A.Kh., Fido, R., Tatham, A.S., and Shewry, P.R. 1996. Disulphide structure of a sunflower seed albumin: Conserved and variant disulphide bonds in the cereal prolamin superfamily. *FEBS Letts.*, **396**:285–288.
- Eigenmann, P.A., Burks, A.W., Bannon, G.A., and Sampson, H.A. 1996. Identification of unique peanut and soy allergens in sera adsorbed with cross-reacting antibodies. *J. Allergy Clin. Immunol.*, **98**:969–978.
- Ericson, M.L., Rodin, J., Lenman, M., Glimelius, K., Josefsson, L.-G., and Rask, L. 1986. Structure of the rapeseed 1.7S storage protein, napin, and its precursor. *J. Biol. Chem.*, **261**:14576–14581.
- Fahlbush, B., Rudeschko, O., Schumann, C., Steurich, R., Henzgen, M., Schlensvoigt, G., and Jager, L. 1998. Further characterisation of IgE-binding antigens in kiwi with particular emphasis on glycoprotein allergens. *J. Investig. Allergol. Clin. Immunol.*, **8**:325–332.
- FAO-WHO. 2001. Evaluation of Allergenicity of Genetically Modified Foods, *Report of the joint FAO/WHO expert consultation on allergenicity of foods derived from Biotechnology*, 22–25th January. pp. 1–26.
- Fischer, J., Becker, C., Hillmer, S., Horstmann, C., Neubohn, B., Schlereth, A., Senyuk, V., Shutov, A. and Müntz, K. 2000. The families of papain- and legumain-like cysteine proteinases from embryonic axes and cotyledons of *Vicia* seeds: Developmental patterns, intracellular localization and functions in globulin proteolysis. *Plant Mol. Biol.*, **43**:83–101.
- Freye, H.B. 1988. Papain anaphylaxis: A case report. *Allergy Proc.*, **9**:571–574.
- Gall, H., Kalveram, K.J., Forck, G., and Sterry, W. 1994. Kiwi fruit allergy—A new birch pollen-associated food allergy. *J. Allergy Clin. Immunol.*, **94**, 70–76.
- Gane, P.J., Dunwell, J.M., and Warwicker, J. 1998. Modeling based on the structure of vicilins predicts a histidine cluster in the active site of oxalate oxidase. *J. Mol. Evol.*, **46**:488–493.
- Garcia-Robaina, J.C., de la Torre-Morin, F., Sanchez-Machin, I., Sanchez-Monge R., Barber, D., and Lombardero, M. 2001. Anaphylaxis induced by exercise and wine. *Allergy*, **56**:357–358.
- Gatehouse, J.A., Croy, R.R.D., and Boulter, D. 1984. The synthesis and structure of pea storage proteins. *Crit. Rev. Plant Sci.*, **1**:287–314.
- Genov, N., Goshev, I., Nikolova, D., Georgieva, D.N., Filippi, B., and Svendsen, I. 1997. A novel thermostable inhibitor of trypsin and subtilisin from the seeds of *Brassica nigra*: Amino acid sequence, inhibitory and spectroscopic properties and thermostability. *Biochim. Biophys. Acta*, **134**:157–164.
- Gianinazzi, S., Martin, C., and Vallée, J.C. 1969. Hypersensibilité aux virus, température et protéines solubles chez le *Nicotiana Xanthi-n.c.* Apparition de nouvelles macromolécules lors de la répression de la synthèse virale. *Compte rendu de l'Académie des Sciences de Paris D*, **268**:800–802.
- Gincel, E., Simorre, J.P., Caille, A., Marion, D., Ptak, M., and Vovelle, F. 1994. Three-dimensional structure in solution of a wheat lipid-transfer protein from multidimensional 1H-NMR data. A new folding for lipid carriers. *Eur. J. Biochem.*, **226**:413–422.
- Gietl C., Wimmer, B., Adamec, J., and Kalousek, F. 1997. A cysteine endopeptidase isolated from castor bean endosperm microbodies processes the glyoxysomal malate dehydrogenase precursor protein. *Plant Physiol.*, **113**:863–871.
- Gomar, J., Sodano, P., Sy, D., Shin, D.H., Lee, J.Y., Suh, S.W., Marion, D., Vovelle, F., and Ptak, M. 1988. Comparison of solution and crystal structures of maize nonspecific lipid transfer protein: A model for a potential in vivo lipid carrier protein. *Proteins-Struct. Funct. Genet.*, **31**:160–171.
- Gomar, J., Petit, M.C., Sodano, P., Sy, D., Marion, D., Kader, J.C., Vovelle, F., and Ptak, M. 1996. Solution structure and lipid binding of a nonspecific lipid transfer protein extracted from maize seeds. *Protein Sci.*, **5**:565–577.
- Gonzales de la Pena, M.A., Monsalve, R.I., Batanero, E., Villalba, M., and Rodriguez, R. 1996. Expression in *Escherichia coli* of Sin a I, the major allergen from yellow mustard seeds. *Eur. J. Biochem.*, **237**:827–832.
- Gourinath, S., Alam, N., Srinivasan, A., Betzel, C., and Singh, T.P. 2000. Structure of the bifunctional inhibitor of trypsin and alpha-amylase from ragi seeds at 2.2 angstrom resolution. *Acta Crystallograph. Sec. D—Biol. Crystallog.*, **56**:287–293.
- Grobe, K., Becker, W.M., Schlaak, M., and Petersen, A. 1999. Grass group I allergens (beta-expansins) are novel, papain-related proteinases. *Eur. J. Biochem.*, **263**:33–40.
- Groves, M.R., Taylor, M.A., Scott, M., Cummings, N.J., Pickersgill, R.W., and Jenkins, J.A. 1996. The prosequence of procaricain forms an alpha-helical domain that prevents access to the substrate-binding cleft. *Structure*, **4**:1193–1203.
- Gueguen, J., Popineau, Y., Anisimova, I.V., Fido, R.J., Shewry, P.R., and Tatham, A.S. 1996. Functionality of the 2S albumin seed storage proteins from sunflower *Helianthus annuus* L. *J. Agric. Food Chem.*, **44**:1184–1189.
- Guénard-Bilbault, L., Kanny, G., and Moneret-Vautrin, D. 1999. L'allergie alimentaire à la farine de blé. *Allergie Immunol.*, **31**:22–25.
- Heinemann, B., Andersen, K.V., Nielsen, P.R., Bech, L.M., and Poulsen, F.M. 1996. Structure in solution of a four-helix lipid binding protein. *Protein Sci.*, **5**:13–23.
- Helm, R.M., Cockrell, G., Herman, E., Burks, A.W., Sampson, H.A., and Bannon, G.A. 1998. Cellular and molecular characterization of a major soybean allergen. *Int. Arch. Allergy Immunol.*, **117**:29–37.
- Helm, R., Cockrell, G., Connaughton, C., Sampson, H.A., Bannon, G.A., Beilinson, V., Livingstone, D., Nielsen, N.C., and Burks, A.W. 2000a. A soybean G2 glycinin allergen. 1. Identification and characterization. *Int. Arch. Allergy Immunol.*, **123**:205–212.
- Helm, R., Cockrell, G., Connaughton, C., Sampson, H.A., Bannon G.A., Beilinson, V., Nielsen, N.C., and Burks, A.W. 2000b. A soybean G2 glycinin allergen. 2. Epitope mapping and three-dimensional modelling. *Int. Arch. Allergy Immunol.*, **123**:213–219.
- Helm, R.M., Cockrell, G., Connaughton, C., West, C.M., Herman, E., Sampson, H.A., Bannon, G.A., and Burks, A.W. 2000c. Mutational analysis of the IgE-binding epitopes of P34/Gly m Bd 30K. *J. Allergy Clin. Immunol.*, **105**:378–384.
- Herian, A.M., Taylor, S.L., and Bush, R.K. 1990. Identification of soybean allergens by immunoblotting with sera from soy-allergic adults. *Int. Arch. Allergy Appl. Immunol.*, **92**:193–198.
- Hewitt, C.R., Brown, A.P., Hart, B.J., and Pritchard, D.I. 1995. A major house dust mite allergen disrupts the immunoglobulin E network by selectively cleaving CD23: Innate protection by antiproteases. *J. Exp. Med.*, **182**:1537–1544.
- Higgins, T.J.V., Chandler, P.M., Randall, P.J., Spencer, D., Beach, L.R., Blagrove, R.J., Kortt, T.J.V., and Inglis, A.S. 1986. Gene structure, protein structure and regulation of the synthesis of a sulfur-rich protein in pea seeds. *J. Biol. Chem.*, **261**:1124–1130.
- Hilder, V.A., Gatehouse, A.M.R., Sheerman, S.E., Barker, R.F., and Boulter, D. 1987. A novel mechanism of insect resistance engineered into tobacco. *Nature*, **330**:160–163.
- Hoffmann-Sommergruber, K. 2000. Plant allergens and pathogenesis-related proteins. *Int. Arch. Allergy Immunol.*, **122**:155–166.
- Hopkin, J.M. (2002). The rise of atopy and links to infection. *Allergy*, **57**(Suppl 72):5–9.

- Hsia, C.C. and Anderson, O.D. 2001. Isolation and characterization of wheat ω -gliadin genes. *Theor. Appl. Genet.*, **103**:37–44.
- Huang, L., Mills, E.N.C., and Morgan, M.R.A. 1998. Analysis of thermal stability of soy globulins using monoclonal antibodies. *Biochim. Biophys. Acta*, **1388**:215–226.
- Irwin, S.D., Keen, J.N., Findlay, J.B.C., and Lord, J.M. 1990. The *Ricinus communis* 2S albumin precursor: A single preprotein may be processed into two different heterodimeric storage proteins. *Mol. Gen. Genet.*, **222**:400–408.
- Ishihara, H., Takahashi, N., Oguri, S., and Tejima, S. (1979). Complete structure of the carbohydrate moiety of stem bromelain. An application of the almond glycopeptidase for structural studies of glycopeptides. *J. Biol. Chem.*, **254**:10715–10719.
- Izumi, H., Adachi, T., Fujii, N., Matsuda, T., Nakamura, R., Tanaka, K., Urisu, A., and Kurosawa, Y. 1992. Nucleotide sequence of a cDNA clone encoding a major allergenic protein in rice seeds. Homology of the deduced amino acid sequence with members of α -amylase/trypsin inhibitor family. *FEBS Letts.*, **301**:213–216.
- Jiang, B.-H., Siregar, U., Willeford, K.O., Luthe, D.S., and Williams, W.P. 1996. Association of a 33-kD cysteine proteinase found in corn callus with the inhibition of fall armyworm larval growth. *Plant Physiol.*, **108**:1631–1640.
- James, J.M., Sixbey, J.P., Helm, R.M., Bannon, G.A., and Burks, A.W. 1997. Wheat α -amylase inhibitor: A second route of allergic sensitisation. *J. Allergy Clin. Immunol.*, **99**:239–244.
- Jones, M.L., Larsen, P.B., and Woodson, W.R. 1995. Ethylene-regulated expression of a carnation cysteine proteinase during flower petal senescence. *Plant Mol. Biol.*, **28**:505–512.
- Kamata Y. and Shibasaki, K. 1978. Formation of digestion intermediate of glycinin. *Agric. Biol. Chem.*, **42**:2323–2329.
- Kalinski, A., Weisemann, J.M., Matthews, B.F., and Herman, E.M. 1990. Molecular cloning of a protein associated with soybean seed oil bodies that is similar to thiol proteases of the papain family. *J. Biol. Chem.*, **265**:13843–13848.
- Kamphuis, I.G., Kalk, K.H., Swarte, M.B., and Drenth, J. 1984. Structure of papain refined at 1.65 Å resolution. *J. Mol. Biol.*, **179**:233–256.
- Kasarda, D.D., Okita, T.W., Bernardin, J.E., Baecker, P.A., Nimmo, C.C., Lew, E.J.-L., Dietler, M.D., and Greene, F. 1984. Nucleic acid (cDNA) and amino acid sequences of α -type gliadins from wheat (*Triticum aestivum*). *Proc. Natl. Academy Sci. USA*, **81**:4712–4716.
- Kelly, J.D., Hlywka, J.J., and Hefle, S.L. 2000. Identification of sunflower seed IgE-binding proteins. *Int. Arch. Allergy Immunol.*, **121**:19–24.
- Kennedy, M.W. 2000. The polyprotein lipid binding proteins of nematodes. *Biochim. Biophys. Acta*, **1476**:149–164.
- Kimber, I., Kervliet, N.I., Taylor, S.L., Astwood, J.D., Sarlo, K., and Dearman, R.J. 1999. Toxicology of protein allergenicity: Prediction and characterization. *Toxicol. Sci.*, **48**:157–162.
- Kleber-Janicke, T., Cramer, R., Appenzeller, U., Schlaak, M., and Becker, W.-M. 1999. Selective cloning of peanut allergens including profilin and 2S albumins, by phage display technology. *Int. Arch. Allergy Immunol.*, **119**:265–274.
- Ko, T.P., Ng, J.D., and McPherson, A. 1993. The three-dimensional structure of canavalin from jack bean (*Canavalia ensiformis*). *Plant Physiol.*, **101**:729–744.
- Ko, T.P., Day, J., and McPherson, A. 2000. The refined structure of canavalin from jack bean in two crystal forms at 2.1 and 2.0 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.*, **56**:411–420.
- Kortt, A.A., Caldwell, J.B., Lilley, G.G., and Higgins, T.J.V. 1991. Amino acid and cDNA sequences of a methionine-rich 2S protein from sunflower seed (*Helianthus annuus* L.). *Eur. J. Biochem.*, **195**:329–334.
- Kreis, M., Forde, B.G., Rahman, S., Mifflin, B.J., and Shewry, P.R. 1985. Molecular evolution of the seed storage proteins of barley, rye and wheat. *J. Mol. Biol.*, **183**:499–502.
- Lane, B.G. 2000. Oxalate oxidases and differentiating surface structure in wheat: Germins. *Biochem. J.*, **349**:309–321.
- Larsen, J.N. and Lowenstein, H. 1996. Allergen Nomenclature. *J. Allergy Clin. Immunol.*, **97**:577–578.
- Lawrence, M.C., Suzuki, E., Varghese, J.N., Davis, P.C., Van Donkelaar, A., Tulloch, P.A., and Colman, P.M. 1990. The three-dimensional structure of the seed storage protein phaseolin at 3 Å resolution. *EMBO J.*, **9**:9–15.
- Lawrence, M.C., Izard, T., Beuchat, M., Blagrove, R.J., and Colman, P.M. 1994. Structure of phaseolin at 2.2 Å resolution. Implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. *J. Mol. Biol.*, **238**:748–776.
- Lee, J.Y., Min, K., Cha, H., Shin, D.H., Hwang, K.Y., and Suh, S.W. 1988. Rice non-specific lipid transfer protein: The 1.6 angstrom crystal structure in the unliganded state reveals a small hydrophobic cavity. *J. Mol. Biol.*, **276**:437–448.
- Leitner, A., Jensen-Jarolim, E., Grimm, R., Wuthrich, B., Ebner H., Scheiner, O., Kraft, D., and Ebner, C. 1998. Allergens in pepper and paprika. Immunologic investigation of the celery-birch-mugwort-spice syndrome. *Allergy*, **53**:36–41.
- Lindorff-Larsen, K., and Winther, J.R. 2001. Surprisingly high stability of lipid transfer protein, LTP1, towards denaturant, heat and proteases. *FEBS Letts.*, **488**:145–148.
- Lilley, G.G. and Inglis, A.S. 1986. Amino acid sequence of conglutin δ , a sulfur-rich seed protein of *Lupinus angustifolius* L. Sequence homology with the C-III α -amylase inhibitor from wheat. *FEBS Letts.*, **195**:235–241.
- Liu, X., Maeda, S., Hu, Z., Aiuchi, T., Nakaya, K., and Kurihara, Y. 1993. Purification, complete amino acid sequence and structural characterization of the heat-stable sweet protein, mabinlin II. *Eur. J. Biochem.*, **211**:281–287.
- Mao, X.-., Sun, D.-J., Miyoshi, A., Feng, Z.T., Handzel, J.M., Hopkin, J.M., and Shirakawa, T. 2000. The link between helminthic infection and atopy. *Parasitol. Today*, **16**:186–188.
- Maleki, S.J., Kopper, R.A., Shin, D.S., Park, C.-W., Compadre, C.M., Sampson, H., Burks, A.W., and Bannon, G.A. 2000. Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation. *J. Immunol.*, **164**:5844–5849.
- Mantjarvi, R., Rautiainen, J., and Virtanen, T. 2000. Lipocalins as allergens. *Biochim. Biophys. Acta*, **1482**:308–317.
- Martin, S.J. and Green, D.R. 1995. Protease activation during apoptosis: Death by a thousand cuts? *Cell*, **82**:349–352.
- Maruyama, N., Ichise, K., Katsube, T., Kishimoto T., Kawase S.-I., Matsumura Y., Takeuchi, Y., Sawada, T., and Utsumi, S. 1998. Identification of major wheat allergens by means of the *Escherichia coli* expression system. *Eur. J. Biochem.*, **255**:739–745.
- Menendez-Arias, L., Moneo, I., Dominguez, J., and Rodriguez, R. 1988. Primary structure of the major allergen of yellow mustard (*Sinapis alba* L.) seed, Sin a I. *Eur. J. Biochem.*, **177**:159–166.
- Mikonen, A., Porali, I., Cercos, M., and Ho, T.-H.D. 1996. A major cysteine proteinase, EPB, in germinating barley seeds: Structure of two intronless genes and regulation of expression. *Plant Mol. Biol.*, **31**:239–254.
- Mills, E.N.C., Huang, L., Noel, T.R., Gunning, A.P. and Morris, V.J. 2001. Formation of thermally-induced aggregates of the soy globulin β -conglycinin. *Biochim. Biophys. Acta*, **1547**:339–350.
- Möller, M., Pascke, A., Vieluf, D., Kayma, M., Vieths, S., and Steinhart, H. 1997. Characterization of allergens in kiwi fruit and detection of cross-reactivities with allergens of birch pollen and related fruits. *Food Agric. Immunol.*, **9**:107–121.
- Monsalve, R.I., Gonzalez de la Pena, M.A., Menezes-Arias, L., Lopez-Otin, C. Villalba, M., and Rodriguez, R. 1993. Characterization of a new oriental-mustard (*Brassica-juncea*) allergen Bra j I: Detection of an allergenic epitope. *Biochem. J.*, **293**:625–632.
- Moneret-Vautrin, D.A., Rance, F., Kanny, G., Olsewski, A., Gueant, J.L., Dutau, G., and Guerin, L. 1998. Food allergy to peanuts in France—evaluation of 142 observations. *Clin. Exp. Allergy*, **28**:1113–1119.
- Moore, J., McDermott, L., Price, N.C., Kelly, S.M., Cooper, A., and Kennedy, M.W. 1999. Sequence-divergent units of the ABA-1 polyprotein array of the nematode *Ascaris suum* have similar fatty-acid- and retinol-binding properties but different binding-site environments. *Biochem. J.*, **340**:337–343.
- Moroz, L.A. and Yang, W.H. 1980. Kunitz soybean trypsin inhibitor; a specific allergen in food anaphylaxis. *N. Engl. J. Med.*, **302**:1126–1128.

- Müntz, K., Jung, R., and Saalbach, G. 1993. Synthesis, processing and targeting of legume seed proteins. In: *Seed Storage Compounds: Biosynthesis, Interactions, and Manipulations*. pp. 128–146. Eds., Shewry, P.R. and Stobart, K., Oxford University Press, Oxford.
- Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.*, **247**:536–540.
- Nakamura, R., and Matsuda, T. 1996. Rice allergenic protein and molecular-genetic approach for hypoallergenic rice. *Biosci. Biotech. Biochem.*, **60**:1215–1221.
- Nakase, M., Adachi, T., Urisu, A., Miyashita, T., Alvarez, A.M., Nagasaka, S., Aoki, N., Nakamura, R., and Matsuda, T. 1996. Rice (*Oryza sativa* L.) α -amylase inhibitors of 14–16 kDa are potential allergens and products of a multigene family. *J. Agric. Food Chem.*, **44**:2624–2628.
- Nielsen, N.C., Dickinson, C.D., Cho, T.J., Thanh, V.H., Scallan, B.J., Fischer, R.L., Sims, T.L., Drews, G.N., and Goldberg, R.B. 1989. Characterization of the glycinin gene family in soybean. *The Plant Cell*, **1**:313–328.
- Nishiyama, I. and Oota, T. 2002. Varietal difference in actinidin concentration and protease activity in the kiwi fruit juice. *J. Jap. Soc. Food Sci. Technol.*, **49**:401–408.
- Nirasawa, S., Liu, C., Nishino, T., and Kurihara, Y. 1993. Disulfide bridge structure of the heat-stable sweet protein mabinlin II. *Biochim. Biophys. Acta.*, **1202**:277–280.
- Nordlee, J.A., Taylor, S.L., Townsend, J.A., Thomas, L.A., and Bush, R.K. 1996. Identification of a Brazil-nut allergen in transgenic soybeans. *N. Engl. J. Med.*, **334**:688–692.
- Oda, Y., Matsunaga, T., Fukuyama, K., Miyazaki, T., and Morimoto, T. 1997. Tertiary and quaternary structures of 0.19 α -amylase inhibitor from wheat kernel determined by X-ray analysis at 2.06 angstrom resolution. *Biochemistry*, **36**:13503–13511.
- Ogawa, T., Tsuji, H., Bando, N., Kitamura, K., Zhu, Y.-L., Hirano, H., and Nishikawa, K. 1993. Identification of the soybean allergenic proteins, Gly m Bd 30K, with the soybean seed 34kDa oil-body-associated protein. *Biosci. Biotech. Biochem.*, **57**:1030–1033.
- Onaderra, M., Monsalve, R.I., Mancheno, J.M., Villalba, M., Martinez del Pozo, A., Gavilanes, J.G., and Rodriguez, R. 1994. Food mustard allergen interaction with phospholipid vesicles. *Eur. J. Biochem.*, **225**:609–615.
- Oommen, A., Kelly, J., Benson, A., and Hefle, S. 2000. Identification of IgE-binding epitopes of the Brazil nut 2S albumin allergen. *J. Allergy Clin. Immunol.*, **105**:402.
- Osborne, T.B. 1924. *The Vegetable Proteins*. 2nd ed. Longmans Green and Co.: London.
- Palosuo, K., Alenius, H., Varjonen, E., Koivuluhta, M., Mikkola, J., Keskinen, H., Kalkkinen, N., and Reunala, T. 1999. A novel wheat gliadin as a cause of exercise-induced anaphylaxis. *J. Allergy Clin. Immunol.*, **103**:912–917.
- Park, J.W., Kang, D.B., Kim, C.W., Ko, S.H., Yum, H.Y., Kim, K.E., Hong, C.S., and Lee, K.Y. 2000. Identification and characterization of the major allergens of buckwheat. *Allergy*, **55**:1035–1041.
- Pastorello, E.A., Conti, A., Pravettoni, V., Farioli, L., Rivolta, F., Ansaloni, R., Ispano, M., Incorvaia, C., Giuffrida, M.G., and Ortolani, C. 1998a. Identification of actinidin as the major allergen of kiwi fruit. *J. Allergy Clin. Immunol.*, **101**:531–537.
- Pastorello, E.A., D'Ambrosio, F.P., Pravettoni, V., Farioli, L., Giuffrida, G., Monza, M., Ansaloni, R., Fortunato, D., Scibola, E., Rivolta, F., Incorvaia, C., Bengtsson, A., Conti, A., and Ortolani, C. 2000a. Evidence for a lipid transfer protein as the major allergen of apricot. *J. Allergy Clin. Immunol.*, **105**:371–377.
- Pastorello, E.A., Farioli, L., Pravettoni, V., Ispano, M., Conti, A., Ansaloni, R., Rotondo, F., Incorvaia, C., Bengtsson, A., Rivolta, F., Trambaioli, C., Previti, M., and Ortolani, C. 1998b. Sensitization to the major allergen of Brazil nut is correlated with the clinical expression of allergy. *J. Allergy Clin. Immunol.*, **102**:1021–1027.
- Pastorello, E.A., Farioli, L., Pravettoni, V., Ispano, M., Scibola, E., Trambaioli, C., Giuffrida, M.G., Ansaloni, R., Godovac-Zimmermann, J., Conti, A., Fortunato, D., and Ortolani, C. 2000b. The maize major allergen, which is responsible for food-induced allergic reactions, is a lipid transfer protein. *J. Allergy Clin. Immunol.*, **106**:744–751.
- Pastorello, E.A., Farioli, L., Pravettoni, V., Ortolani, C., Ispano, M., Monza, M., Baroglio, C., Scibola, E., Ansaloni, R., Incorvaia, C., and Conti, A. 1999a. The major allergen of peach (*Prunus persica*) is a lipid transfer protein. *J. Allergy Clin. Immunol.*, **103**:520–526.
- Pastorello, E.A., Farioli, L., Robino, A.M., Trambaioli, C., Conti, A., and Pravettoni, V. 2001a. A lipid transfer protein involved in occupational sensitization to spelt. *J. Allergy Clin. Immunol.*, **108**:145–146.
- Pastorello, E.A., Ortolani, C., Baroglio, C., Pravettoni, V., Ispano, M., Giuffrida, M.G., Fortunato, D., Farioli, L., Monza, M., Napolitano, L., Sacco, M., Scibola, E., and Conti, A. 1999b. Complete amino acid sequence determination of the major allergen of peach (*Prunus persica*) Pru p 1. *J. Biol. Chem.*, **380**:1315–1320.
- Pastorello, E.A., Pravettoni, V., Farioli, L., Ispano, M., Fortunato, D., Monza, M., Giuffrida, M.G., Rivolta, F., Scibola, E., Ansaloni, R., Incorvaia, C., Conti, A., and Ortolani, C. 1999c. Clinical role of a lipid transfer protein that acts as a new apple-specific allergen. *J. Allergy Clin. Immunol.*, **104**:1099–1106.
- Pastorello, E.A., Pravettoni, V., Ispano, M., Farioli, L., Ansaloni, R., Rotondo, F., Incorvaia, C., Asman, I., Bengtsson, A., and Ortolani, C. 1996. Identification of the allergenic components of kiwi fruit and evaluation of their cross-reactivity with timothy and birch pollens. *J. Allergy Clin. Immunol.*, **98**:601–610.
- Pastorello, E.A., Varin, E., Farioli, L., Pravettoni, V., Ortolani, C., Trambaioli, C., Fortunato, D., Giuffrida, M.G., Rivolta, F., Robino, A., Calamari, A.M., Lacava, L., and Conti, A. 2001b. The major allergen of sesame seeds (*Sesamum indicum*) is a 2S albumin. *J. Chromatogr. B Biomed. Sci. Appl.*, **756**:85–93.
- Pauli, G., Oster, J.P., Deviller, P., Heiss, S., Bessot, J.C., Susani, M., Ferreira, F., Kraft, D., and Valenta, R. 1996. Skin testing with recombinant allergens rBet v 1 and birch profilin, rBet v 2: Diagnostic value for birch pollen and associated allergies. *J. Allergy Clin. Immunol.*, **97**:1100–1109.
- Pechan, T., Ye, L., Chang, Y.-M., Mitra, A., Lin, L., Davis, F.M., Williams, W.P., and Luthe, D.S. 2000. A unique 22-kD cysteine proteinase accumulates in response to larval feeding in maize genotypes resistant to fall armyworm and other lepidoptera. *The Plant Cell*, **12**:1031–1040.
- Pedrosa, C., De Felice, F.G., Trisciuzzi, C., and Ferreira, S.T. 2000. Selective neoglycosylation increases the structural stability of vicilin, the 7S storage protein of pea seeds. *Arch. Biochem. Biophys.*, **15**:203–210.
- Plumb, G.W., Mills, E.N.C., Tatton, M.J., D'Ursel, C.C.M., Lambert, N., and Morgan, M.R.A. 1994. The effect of thermal and proteolytic processing on glycinin, the 11S globulin of soy (*Glycine max*)—a study utilising monoclonal and polyclonal antibodies. *J. Agric. Food Chem.*, **42**:705–710.
- Poltronieri, P., Cappello, M.S., Dohmae, N., Conti, A., Fortunato, D., Pastorello, E.A., Ortolani, C., and Zacheo, G. 2002. Identification and characterisation of the IgE-binding proteins 2S albumin and conglutin- γ in almond (*Prunus dulcis*) seeds. *Int. Arch. Allergy Immunol.*, **128**:97–104.
- Polya, G.M., Chandra, S., and Condrón, R. 1993. Purification and sequencing of radish seed calmodulin antagonists phosphorylated by calcium-dependent protein kinase. *Plant Physiol.*, **101**:545–551.
- Poznanski, J., Sodano, P., Suh, S.W., Lee, J.Y., Ptak, M., and Vovelle, F. 1999. Solution structure of a lipid transfer protein extracted from rice seeds—Comparison with homologous proteins. *Euro. J. Biochem.*, **259**:692–708.
- Rabjohn, P., Helm, E.M., Stanley, J.S., West, C.M., Sampson, H.A., Burks, A.W., and Bannon, G.A. 1999. Molecular cloning and epitope analysis of the peanut allergen Ara h 3. *J. Clin. Invest.*, **103**:535–542.
- Rawlings, N.D. and Barrett, A.J. 1993. Evolutionary families of peptidases. *Biochem. J.*, **290**:205–218.
- Richardson, M. 1991. Seed storage proteins: The enzyme inhibitors. *Methods in Plant Biochemistry*, **5**:259–305.
- Rico, M., Bruix, M., Gonzalez, C., Monsalve, R.I., and Rodriguez, R. 1996. H-1 NMR assignment and global fold of napin BnIb, a representative 2S albumin seed protein. *Biochem.*, **35**:5672–5682.
- Roux, K.H., Teuber, S.S., Robotham, J.M., and Sathe, S.K. 2001. Detection and stability of the major almond allergen in foods. *J. Agric. Food Chem.*, **49**:2131–2136.

- Salvi, S. 2001. Pollution and allergic airways disease. *Curr. Opin. Allergy Clin. Immunol.*, **1**:35–41.
- Samoto, M., Miyazaki, C., Akasaka, T., Mori, H., and Kawamura, Y. 1996. Specific binding of allergenic soybean protein Gly m Bd 30K with α - and α -subunits of conglycinin in soy milk. *Biosci. Biotech. Biochem.*, **60**:1006–1010.
- Sanchez-Monge, R., Lombardero, M., Garcia-Selles, F.J., Barber, D., and Salcedo, G. 1999. Lipid-transfer proteins are relevant allergens in fruit allergy. *J. Allergy Clin. Immunol.*, **103**: 514–519.
- Sanchez-Monge, R., Pascual, C.Y., Diaz-Perales, A., Fernandez-Crespo, J., Martin-Esteban, M., and Salcedo, G. 2000. Isolation and characterisation of relevant allergens from boiled lentils. *J. Allergy Clin. Immunol.*, **106**:955–961.
- Scheurer, S., Pastorello, E.A., Wangorsch, A., Kastner, M., Haustein, D., and Vieths, S. 2001. Recombinant allergens Pru av 1 and Pru av 4 and a newly identified lipid transfer protein in the in vitro diagnosis of cherry allergy. *J. Allergy Clin. Immunol.*, **107**:724–731.
- Schulz, O., Laing, P., Sewell, H.F., and Shakib, F. 1995. Der p I, a major allergen of the house dust mite, proteolytically cleaves the low-affinity receptor for human IgE (CD23). *Eur. J. Immunol.*, **25**:3191–3194.
- Seppala, U., Majamaa, H., Turjanmaa, K., Helin, J., Reunala, T., Kalkkinen, N., and Palosuo, T. 2001. Identification of four novel potato (*Solanum tuberosum*) allergens belonging to the family of soybean trypsin inhibitors. *Allergy*, **56**:619–626.
- Shewry, P.R. 1999. Enzyme inhibitors of seeds: Types and properties. In: *Seed Proteins*. pp. 587–615. Eds., Shewry, P.R., and Casey, R., Kluwer Academic Publishers: Dordrecht, The Netherlands.
- Shewry, P.R. and Tatham, A.S. 1990. The prolamin storage proteins of cereal seeds: Structure and evolution. *Biochem. J.*, **267**:1–12.
- Shewry, P.R. and Casey, R. Eds. 1999. *Seed Proteins*. Kluwer Academic Publishers: Dordrecht, The Netherlands.
- Shewry, P.R. and Lucas, J.A. 1997. Plant proteins that confer resistance to pests and pathogens. In: *Advances in Botanical Research*. Vol. 26, pp. 135–192. Ed., Callow, J., Academic Press.
- Shewry, P.R. and Pandya, M.J. 1999. The 2S albumin storage proteins. In: *Seed Proteins* pp. 563–586. Eds., Shewry, P.R., and Casey, R., Kluwer Academic Publishers, Dordrecht.
- Shewry, P.R. and Tatham, A.S. 1999. The characterisation, structures and evolutionary relationships of prolamins. In: *Seed Proteins* pp. 11–33. Eds., Shewry, P.R., and Casey, R., Kluwer Academic Publishers: Dordrecht.
- Shewry, P.R., Tatham, A.S., and Halford, N.G. 1999. The prolamins of the triticeae. In: *Seed Proteins* pp. 35–78. Eds., Shewry, P.R. and Casey, R., Kluwer Academic Publishers, Dordrecht.
- Shibasaki, M., Suzuki, S., Tajima, S., Nemoto, H., and Kuroume, T. 1980. Allergenicity of major component proteins of soybean. *Int. Arch. Allergy Appl. Immunol.*, **61**:441–448.
- Shin, D.H., Lee, J.Y., Hwang, K.Y., Kim, K.K., and Suh, S.W. 1995. High-resolution crystal-structure of the nonspecific lipid-transfer protein from maize seedlings. *Structure*, **3**:189–199.
- Shin, D.S., Compadre, C.M., Maleki, S.J., Kopper, R.A., Sampson, H., Huang, S.K., Burks, A.W., and Bannon, G.A. 1998. Biochemical and structural analysis of the IgE binding sites on Ara h 1, an abundant and highly allergenic peanut protein. *J. Biol. Chem.*, **273**:13753–13759.
- Solomon, M., Belenghi, B., Delledonne, M., Menachem, E., and Levine, A. 1999. The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. *Plant Cell*, **11**:431–444.
- Soto-Mera, M.T., Lopez-Rico, M.R., Filgueira, J.F., Villamil, E., and Cidras, R. 2000. Occupational allergy to papain. *Allergy*, **55**:983–984.
- Stewart, G.A., and Thompson, P.J. 1996. The biochemistry of common aeroallergens. *Clin. Exp. Allergy*, **26**:1020–1044.
- Stanley, J.S., King, N., Burks, A.W., Huang, S.K., Sampson, H., Cockrell, G., Helm, R.M., West, C.M., and Bannon, G.A. 1997. Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Ara h 2. *Arch. Biochem. Biophys.*, **342**:244–253.
- Strobl, S., Mühlhahn, P., Bernstein, R., Wilschek, R., Maskos, K., Wunderlich, M., Huber, R., Glockhuber, R., and Holak, T.A. 1995. Determination of the three-dimensional structure of the bifunctional α -amylase/trypsin inhibitor from ragi seeds by NMR spectroscopy. *Biochemistry*, **34**:8281–8293.
- Sumner, I.G., Harris, G.W., Taylor, M.A., Pickersgill, R.W., Owen, A.J., and Goodenough, P.W. 1993. Factors effecting the thermostability of cysteine proteinases from *Carica papaya*. *Eur. J. Biochem.*, **214**:129–134.
- Svendsen, I., Nicolova, D., Goshev, I., and Genov, N. 1989. Isolation and characterisation of a trypsin inhibitor from the seeds of khorabi (*Brassica napus* var *Rapifera*) belonging to the napin family of storage proteins. *Carlsberg Res. Commun.*, **54**:231–239.
- Svendsen, I., Nicolova, D., Goshev, I., and Genov, N. 1994. Primary structure, spectroscopic and inhibitory properties of a two-chain trypsin inhibitor from the seeds of charlock (*Sinapis arvensis* L.), a member of the napin protein family. *Int. J. Peptide Protein Res.*, **43**:425–430.
- Takizawa, T., Arakawa, H., Tokuyama, K., and Morikawa, A. 2001. Identification of allergen fractions of wheat flour responsible for anaphylactic reactions to wheat products in infants and young children. *Int. Arch. Allergy Immunol.*, **125**:51–56.
- Tanabe, S., Arai, S., Yanagihara, Y., Takahashi, K., and Watanabe, M. 1996. A major wheat allergen has a Gln-Gln-Gln-Pro-Pro motif identified as an IgE-binding epitope. *Biochem. Biophys. Res Commun.*, **219**:290–293.
- Taneva, S.G., Donchev, A.A., Dimitrov, M.I., and Muga, A. 2000. Redox- and pH-dependent association of plastocyanin with lipid bilayers: Effects on protein conformation and thermal stability. *Biochim. Biophys. Acta*, **1463**:429–438.
- Tassin-Moindrot, S., Caille, A., Douliez, J.-P., Marion, D., and Vovelle, F. 2000. The wide binding properties of a wheat nonspecific lipid transfer protein—Solution structure of a complex with prostaglandin B-2. *Euro. J. Biochem.*, **267**:1117–1124.
- Tejera, M.L., Villalba, M., Batanero, E., and Rodriguez, R. 1999. Identification, isolation, and characterization of Ole e 7, a new allergen of olive tree pollen. *J. Allergy Clin. Immunol.*, **104**:797–802.
- Terras, F.R.G., Schoofs, H.M., Thevissen, K., Osborn, R.W., Vanderleyden, J., Cammue, B.P.A., and Broekaert, W.F. 1993. Synergistic enhancement of the antifungal activity of wheat and barley thionins by radish and oilseed rape 2S albumins and by barley trypsin inhibitors. *Plant Physiol.*, **103**:1311–1319.
- Terras, F.R.G., Schoofs, H.M., De Bolle, M.F.C., Van Leuven, F., Rees, S.B., Vanderleyden, J., Cammue, B.P.A., and Broekaert, W.F. 1992. Analysis of two novel classes of plant anti-fungal proteins from radish (*Raphanus sativus* L.) seeds. *J. Biol. Chem.*, **267**:15301–15309.
- Teuber, S.S., Dandekar, A.M., Peterson, W.R., and Sellers, C.L. 1998. Cloning and sequencing of a gene encoding a 2S albumin seed storage protein precursor from English walnut (*Juglans regia*), a major food allergen. *J. Allergy Clin. Immunol.*, **101**:807–814.
- Teuber, S.S., and Petersen, W.R. 1999. Systemic allergic reaction to coconut (*Cocos nucifera*) in 2 subjects with hypersensitivity to tree nut and demonstration of cross-reactivity to legumin-like seed storage proteins: New coconut and walnut food allergens. *J. Allergy Clin. Immunol.*, **103**:1180–1185.
- Teuber, S., Jarvis, K.C., Dandekar, A.M., Peterson, W.R., and Ansari, A. 1999. Identification and cloning of a complementary DNA encoding a vicilin-like proprotein, Jug r 2, from English walnut kernel (*Juglans regia*), a major food allergen. *J. Allergy Clin. Immunol.*, **104**:1111–1120.
- Thanh, V.H. and Shibasaki, K. 1979. Major proteins of soybean seeds. Reversible and irreversible dissociation of β -conglycinin. *J. Agric. Food Chem.*, **27**:805–809.
- Thoyts, P.J.E., Napier, J.A., Millichip, M., Stobart, A.K., Griffiths, W.T., Tatham, A.S., and Shewry, P.R. 1996. Characterization of a sunflower seed albumin which associates with oil bodies. *Plant Sci.*, **118**:119–125.
- Tsuboi, S., Suga, T., Takishima, K., Mamiya, G., Matsui, K., Ozeki, Y., and Yamada, M. 1991. Organ-specific occurrence and expression of the isoforms of nonspecific lipid transfer protein in castor bean seedlings, and molecular-cloning of a full-length cDNA for a cotyledon-specific isoform. *JBiochem.*, **110**:823–831.
- van Loon, L.C. and van Strien, E.A. 1999. The families of pathogenesis-related proteins, their activities and comparative analysis of PR-1 type proteins. *Physiol. Mol. Plant Pathol.*, **55**:85–97.

- van Ree, R., Cabanes-Macheteau, M., Akkerdaas, J., Milazzo, J.P., Loutelier-Bourhis, C., Rayon, C., Villalba, M., Koppelman, S., Aalberse, R., Rodriguez, R., Faye, L., and Lerouge, P. 2000. Beta(1,2)-xylose and alpha(1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J. Biol. Chem.*, **275**:11451–11458.
- Varjonen, E., Vainio, E., and Kalimo, K. 2000. Antigliadin IgE-indicator of wheat allergy in atopic dermatitis. *Allergy*, **55**:386–391.
- Varjonen, E., Vainio, E., and Kalimo, K. 1997. Life-threatening, recurrent anaphylaxis caused by allergy to gliadin and exercise. *Clin. Exp. Allergy*, **27**:162–166.
- Varjonen, E., Vainio, E., Kalimo, K., Juntunen-Backman, K., and Savolainen, J. 1995. Skin-prick test and RAST responses to cereals in children with atopic dermatitis. Characterization of IgE-binding components in wheat and oats by an immunoblotting method. *Clin. Exp. Allergy*, **25**:1100–1107.
- Vioque, J., Sanchez-Vioque, R., Clemente, A., Pedroche, J., Bautista, J., and Millan, F. 1999. Purification and partial characterization of chickpea 2S albumin. *J. Agric. Food Chem.*, **47**:1405–1409.
- Wal, J.-M. 1998. Cow's milk allergens. *Allergy*, **53**:1013–1022.
- Wang, F., Robotham, J.M., Teuber, S.S., Tawde, P., Sathe, S.K., and Roux, K.H. 2002. Ana o 1, a cashew (*Anacardium occidentale*) allergen of the vicilin seed storage protein family. *J. Allergy Clin. Immunol.*, **110**:160–166.
- Woo, E.J., Dunwell, J.M., Goodenough, P.W., Marvier, A.C., and Pickersgill, R.W. 2000. Germin is a manganese containing homohexamer with oxalate oxidase and superoxide dismutase activities. *Nat. Struct. Biol.*, **7**:1036–1040.
- Wright, D.J. 1987. The seed globulins. **In:** *Developments in Food Proteins*. pp. 81–157. Ed., Hudson, B.J.F. Elsevier Applied Science, London and New York.
- Wright, D.J. 1988. The seed globulins—part II. **In:** *Developments in Food Proteins*. pp. 119–178. Ed., Hudson, B.J.F. Elsevier, London.
- Yaklich, R.W., Helm, R.M., Cockrell, G., and Herman, E.M. 1999. Analysis of the distribution of the major soybean seed allergen in a core collection of *Glycine max* accessions. *Crop Sci.*, **39**:1444–1447.
- Yamauchi, F., Yamagishi, T., and Iwabuchi, S. 1991. Molecular understanding of heat-induced phenomena of soybean protein. *Food Rev. Internat.*, **7**:283–322.
- Youle, R.J. and Huang, A.H.C. 1981. Occurrence of low molecular weight and high cysteine containing albumin storage proteins in oilseeds of disperse species. *Am. J. Botany*, **68**:44–48.
- Zavodszky, M., Chen, C.W., Huang, J.K., Zolkiewski, M., Wen, L., Krishnamoorthi, R. 2001. Disulfide bond effects on protein stability: Designed variants of Cucurbita maxima trypsin inhibitor-V. *Protein Sci.*, **10**:149–160.
- Zentner, A., Jeep, S., Wahl, R., Kunkel, G., and Kleine-Tebbe, J. 1997. Multiple IgE-mediated sensitizations to enzymes after occupational exposure: Evaluation by skin prick test, RAST, and immunoblot. *Allergy*, **52**:928–934.