7

Plant Lectins: Versatile Proteins with Important Perspectives in Biotechnology

WILLY J. PEUMANS* AND ELS J.M. VAN DAMME

Katholieke Universiteit Leuven, Laboratory for Phytopathology and Plant Protection, Willem de Croylaan 42, 3001 Leuven, Belgium

Introduction

Many wild and cultivated plants accumulate so-called 'lectins' or agglutinins in their seeds or vegetative tissues. For the sake of simplicity lectins can be defined as (glyco)proteins which bind reversibly to specific mono- or oligosaccharides without altering the structure of the bound ligand. The initial discovery of lectins dates back to 1888 when Stillmark found a proteinaceous haemagglutinating factor in castor beans (*Ricinus communis*). Meanwhile, several hundred lectins/haemagglutinins/phytohaemagglutinins have been isolated and studied in some detail at the biochemical and/or physicochemical level. In addition, many lectin genes have been cloned and the three-dimensional structure of over a dozen lectins has been resolved by X-ray diffraction studies.

The broad interest in plant lectins is undoubtedly related to their highly specific interaction with carbohydrates and the biological effects based thereon. The large impact of plant lectins on biological and biomedical research can be explained for two reasons. First, plant lectins are a readily accessible source of carbohydrate binding proteins. Secondly, plant lectins are particularly suited for the analysis and isolation of animal and human glycoconjugates because the latter compounds are the natural targets of most carbohydrate-binding proteins present in plants. Plant lectins have classically been used as analytical or preparative tools in glycoconjugate research, and as bioactive proteins for the induction of some particular processes in cells or organisms. The recent progress in plant lectin research, especially with respect to the understanding of the structure/specificity/function relationships of the different lectin groups certainly will refine and extend these applications. More importantly, the introduction of modern biotechnology now creates novel opportunities to maximally exploit the unique biochemical and biological properties of plant lectins for previously unconceived applications. This contribution aims to give a critical overview of the perspectives plant lectins can offer in tomorrow's biotechnology. Before discussing

*To whom correspondence may be addressed.

Biotechnology and Genetic Engineering Reviews – Vol. 15, April 1998
0264-8725/98/15/199-228 $20.00 + $0.00

the potential applications of lectins, a brief summary is given of the state of the art of the current knowledge of those aspects of plant lectins which are relevant in view of their use in biotechnology. First, evidence is presented that the apparently very heterogeneous group of lectins can be subdivided into a small number of evolutionary and structurally related families of proteins. Secondly, the carbohydrate-binding specificity of lectins is critically reviewed whereby the emphasis is put on their obvious preference for foreign glycans. Thirdly, the specificity of plant lectins is related to their physiological role and especially to their involvement in the plant's defence. Fourthly, the impact of the occurrence of lectins in crop plants on the safety and quality of the food and feed products derived thereof is discussed.

Carbohydrate-binding proteins from plants: an old field revisited

PLANT LECTINS: A COMPOSITE OF FAMILIES OF STRUCTURALLY AND EVOLUTIONARY RELATED PROTEINS

Before breakthrough of molecular biology, plant lectins were usually regarded as a very heterogeneous group of proteins, differing from each other with respect to their molecular structure, carbohydrate-binding specificity and biological activities. Although the previously observed biochemical/physicochemical heterogeneity obviously reflects a biological reality, sequencing and molecular cloning of lectins from many plant species now allows a fairly simple classification system to be elaborated. Based on the available sequence data, plant lectins can be divided into a relatively small number of natural families of evolutionarily related proteins. Prior to discussion of the different lectin subgroups, some important issues have to be addressed with regard to the definition of the term 'plant lectin'. Because of some novel insights into the structure of lectins and lectin genes the definition of plant lectins has recently been updated. According to the new definition *'all plant proteins possessing at least one non-catalytic domain, that binds reversibly to a specific mono- or oligosaccharide'* are considered as lectins (Peumans and Van Damme, 1995). Since a broad range of proteins with different agglutination and/or glycoconjugate precipitation properties comply with this definition lectins are further subdivided in merolectins, hololectins, chimerolectins and superlectins according to their overall structure. Merolectins consist exclusively of a single carbohydrate-binding domain (e.g. hevein, a chitin-binding latex protein from the rubber tree *Hevea brasiliensis*). Due to their monovalent character merolectins are incapable of precipitating glycoconjugates or agglutinating cells. Hololectins also consist exclusively of carbohydrate-binding domains but contain at least two such domains which are identical or very similar. Because hololectins are di- or multivalent they can agglutinate cells and/or precipitate glycoconjugates. Most, but not all, agglutinating plant lectins are hololectins. Superlectins are composed of at least two carbohydrate-binding domains. Unlike hololectins the carbohydrate-binding domains of the superlectins are not identical or similar but recognize structurally different sugars. For example, the tulip bulb lectin TxLCI which is built up of two dissimilar tandem arrayed carbohydrate-binding domains recognizing mannose and GalNAc, respectively, is a superlectin (Van Damme *et al.*, 1996d). Chimerolectins are 'fusion' proteins built up of a carbohydrate-binding domain tandemly arrayed with an unrelated domain. The latter domain has a

well-defined catalytic (or another biological) activity that acts independently of the carbohydrate-binding domain. Chimerolectins behave as merolectins or hololectins depending on the number of sugar-binding sites. Multivalent type 2 ribosome-inactivating proteins (RIP) (Barbieri *et al.*, 1993), for instance, readily agglutinate cells whereas monovalent class I plant chitinases (with a single chitin-binding domain) (Collinge *et al.*, 1993) do not.

The vast majority of all currently known plant lectins can be classified into a relatively small number of families of evolutionary and structurally related proteins. Seven lectin families are distinguished namely (i) the legume lectins, (ii) the monocot mannose-binding lectins, (iii) the chitin-binding proteins containing hevein domain(s), (iv) the type 2 ribosome-inactivating proteins, (v) the Cucurbitaceae phloem lectins, (vi) the jacalin family and (vii) the Amaranthaceae lectins (*Table 1*).

Table 1. Families of plant lectins: occurrence, molecular structure and specificity

Lectin group	Occurrence		Molecular structure		Nominal specificity
	Taxonomic distribution	Number of identified lectins	Protomer size (kDa)	Number of protomers	
Legume lectins	Legumes	>100	30 or [(30–X)+X][a]	2 or 4	Diverse
Monocot mannose-binding lectins	Liliales	>50	12	1, 2, 3 or 4	Mannose
Chitin-binding lectins	Monocots and dicots	>100	5–35	1 or 2	GlcNAc or $(GlcNAc)_n$
Type 2 RIP	Monocots and dicots	>20	[30-s-s-35][b]	1, 2 or 4	Gal, GalNAc or Neu5Acα(2,6)Gal/GalNAc
Cucurbitaceae phloem lectins	Cucurbitaceae	<10	24	2	$(GlcNAc)_n$
Jacalin family	Moraceae Convolvulaceae	<10	16 or [(16–X)+X][a]	2	Gal Mannose/Maltose
Amaranthaceae lectins	Amaranthaceae	<10	30	2	GalNAc

[a] Protomers are cleaved into two subunits
[b] Both subunits are linked through a disulphide bridge

Legume lectins

Legume lectins are the best known lectin family (Sharon and Lis, 1990). Until now the classical legume lectins have been found exclusively in members of the Leguminoseae. Protein and gene sequencing has demonstrated that all legume lectins are built up of either two or four protomers of about 30 kDa. In the so-called two chain legume lectins the protomers are further cleaved into two smaller polypeptides. All legume lectins are undoubtedly related at the molecular level but show a remarkable variation in carbohydrate-binding specificity. This broad range of specificities certainly contributed to the success of legume lectins as tools. In addition, legume lectins received a lot of attention because of their abundance in many crop plants and their presumed involvement in the symbiosis between legumes and the nitrogen fixing bacterium *Rhizobium* (Diaz *et al.*, 1989).

Monocot mannose-binding lectins

The monocot mannose-binding proteins are a relatively new group of lectins. Since the initial discovery of a lectin with an exclusive specificity towards mannose in snowdrop bulbs (Van Damme *et al.*, 1987) similar proteins have been found in species of the monocot families Amaryllidaceae, Alliaceae, Araceae, Orchidaceae, Liliaceae and Bromeliaceae (Van Damme *et al.*, 1995b). All monocot mannose-binding proteins are built up of 1, 2, 3 or 4 subunits of about 12 kDa and exhibit an exclusive specificity towards mannose. Sequence comparisons clearly demonstrated that all the monocot mannose-binding proteins belong to a single superfamily of evolutionary related proteins. Molecular modelling studies further indicated that they all have a very similar overall structure (Barre *et al.*, 1996) and that the number of active binding sites varies between one and three per protomer. Due to their exclusive specificity towards mannose several monocot mannose-binding proteins have become important tools. In addition, monocot mannose-binding proteins receive a lot of attention because of their potent antiviral (Balzarini *et al.*, 1991, 1992) and anti-insect properties (Gatehouse *et al.*, 1995).

Chitin-binding proteins composed of hevein domains

Chitin-binding proteins containing so-called hevein domain(s) also represent a large lectin family. The term 'hevein domain' refers to hevein, a small 43 amino-acid residue protein found in the latex of the rubber tree (*Hevea brasiliensis*). Hevein is a merolectin composed of a single chitin-binding domain, which is derived from a larger precursor molecule (Lee *et al.*, 1991). Many plant proteins contain chitin-binding domains similar to hevein. Some of these proteins are built up of a single hevein domain, which in some instances is even truncated, e.g. the chitin-binding antimicrobial peptides from *Amaranthus caudatus* (Broekaert *et al.*, 1992). Other chitin-binding proteins are composed of two, three or four tandemly arrayed hevein domains, like the lectins from stinging nettle, *Phytolacca americana* and wheat, respectively (Raikhel *et al.*, 1993; Konami *et al.*, 1995). Some chitin-binding proteins are built up of one or more hevein domains linked to an unrelated domain. Class I chitinases, for instance, are chimerolectins composed of a single hevein domain linked to a catalytically active chitinase domain (Collinge *et al.*, 1993). Similarly, Solanaceae lectins are built up of three tandemly arrayed hevein domains linked to a glycosylated serine-hydroxyproline-rich domain that resembles the cell wall protein extensin (Kieliszewski *et al.*, 1994; Allen *et al.*, 1996). Several chitin-binding lectins built up of hevein domains are derived from chimeric precursors. For instance hevein and the nettle lectin are formed by the cleavage of an extended C-terminal peptide with high sequence similarity to the catalytic domain of plant chitinases (Lee *et al.*, 1991; Lerner and Raikhel, 1992). Proteins containing hevein domain(s) are ubiquitous in the plant kingdom since most probably all higher plants contain class I chitinases. Genuine (agglutinating) chitin-binding lectins are far less widespread than chitinases but have already been found in the families Gramineae, Urticaceae, Viscaceae, Phytolaccaceae, Papaveraceae and Solanaceae (Raikhel *et al.*, 1993; Konami *et al.*, 1995; Peumans *et al.*, 1996). Since all these lectins exhibit a high degree of sequence similarity, one can reasonably assume that the chitin-binding lectins are a superfamily of evolutionary related proteins occurring both in monocot and dicot species.

Type 2 ribosome-inactivating proteins

Type 2 ribosome-inactivating proteins are chimerolectins composed of a polynucleotide:adenosine glycosidase domain (the so-called A chain) tandemly arrayed with a carbohydrate-binding domain (the so-called B chain) (Barbieri *et al.*, 1993). Both chains are synthesized on a single precursor molecule, which is post-translationally processed through the excision of a linker between the A and B chains. Since the A and B chains remain linked through an interchain disulphide bridge, the building block of type 2 RIP corresponds to an [A-s-s-B]-pair. All known type 2 RIP are composed of either one, two or four [A-s-s-B]-pair(s). Usually, the [A-s-s-B]-pairs are held together by non-covalent interactions. Only in the case of the tetrameric Neu5Acα(2,6)Gal/GalNAc-binding *Sambucus* lectins, the [A-s-s-B]-pairs are pairwise linked through an intermolecular disulphide bridge between two adjacent B chains (Van Damme *et al.*, 1996b; Kaku *et al.*, 1996). All type 2 RIP share a high sequence similarity both in the A and the B chain, and have a similar overall folding and three-dimensional structure. In spite of the obvious similarities type 2 RIP strongly differ from each other with respect to their catalytic activity, substrate specificity and (cyto)toxic properties. Most type 2 RIP preferentially bind either Gal or GalNAc. However, the occurrence of at least two types of Neu5Acα(2,6)Gal/GalNAc-binding type 2 RIP in *Sambucus* species (Van Damme *et al.*, 1996b, 1997b) indicates that there is also some divergence in carbohydrate-binding specificity. Type 2 RIP have been found hitherto in the plant families Euphorbiaceae, Fabaceae, Sambucaceae, Viscaceae, Ranunculaceae, Lauraceae, Passifloraceae, Iridaceae and Liliaceae. There is no doubt, therefore, that type 2 RIP also are a superfamily of evolutionary related proteins occurring in different taxonomic groups. Moreover, the recent discovery of several novel type 2 RIP suggests that they are probably much more widespread in the plant kingdom than is actually believed.

Jacalin family

Jacalin, the galactose-specific lectin from seeds of jackfruit (*Artocarpus integrifolia*) is one of the classical non-legume plant lectins. Similar lectins have also been found in a few other *Artocarpus* species and in *Maclura pomifera* (Young *et al.*, 1989). All Moraceae lectins are built up of two identical protomers. Each protomer contains a small and a large subunit, both derived from a single precursor through a complex post-translational modification of the primary translation product. Until recently, the Moraceae lectins were considered as a small lectin family which occurs exclusively in this particular plant family. However, molecular cloning of Calsepa, a mannose/maltose binding lectin from rhizomes of *Calystegia sepium* (family Convolvulaceae), revealed that this novel lectin shares a remarkable sequence similarity with jacalin (Van Damme *et al.*, 1996a). Calsepa is also composed of two protomers. Unlike in jacalin, the protomers of Calsepa consist of a single polypeptide chain of 153 amino acids. Other members of the Convolvulaceae family like *Convolvulus arvensis* and *Ipomea batatas* also contain lectins similar to Calsepa (W.P., unpublished results). Moreover, there are indications for the occurrence of related lectins in other plant families. Most likely, the jacalin family also is a superfamily of lectins occurring in several taxonomic groups. It is evident, however, that the degree of sequence similarity within this family is lower than in the above mentioned lectin families.

Moreover, the jacalin family is also heterogeneous with respect to the carbohydrate-binding specificity.

Cucurbitaceae phloem lectins

Many Cucurbitaceae species contain high concentrations of a lectin that binds oligomers of GlcNAc (Wang *et al.*, 1994). These so-called Cucurbitaceae phloem lectins are dimeric proteins composed of two identical subunits of about 24 kDa. All currently known Cucurbitaceae phloem lectins show a high degree of sequence similarity but have no apparent sequence similarity with any other plant lectin or protein. Although it cannot be excluded that similar lectins occur in other taxonomic groups, it seems likely that the Cucurbitaceae phloem lectins are a small protein family confined to a relatively small taxonomic group.

Amaranthaceae lectins

Molecular cloning and X-ray diffraction analysis revealed that amaranthin, the seed lectin from *Amaranthus caudatus* does not resemble any other plant lectin with respect to its amino acid sequence (Raina and Datta, 1992) and three-dimensional structure (Transue*et al.*, 1997). Based on these data, amaranthin is now considered as the prototype of the amaranthin lectin family. Several other *Amaranthus* species (e.g. *A. spinosus*, *A. leucocarpus* and *A. cruentus*) contain lectins which are very similar to amaranthin. No related lectins (or other proteins) have been found in any other plant family. Although the latter observation does not preclude the occurrence of amaranthins or amaranthin-related proteins in other species, the amaranthins are, for the time being, regarded as a small lectin family confined to a relatively inconspicuous part of the plant kingdom.

The vast majority of all plant lectins belongs to one of the seven lectin families described above. At present, several plant lectins cannot be classified because of the lack of sequence information. Some of the unclassified lectins probably belong to one of the seven lectin families. It is also possible, however, that novel lectin families will be discovered when more sequence data become available.

THE CARBOHYDRATE-BINDING SPECIFICITY OF PLANT LECTINS

Lectins distinguish themselves from all other plant proteins by their capability to bind simple or complex carbohydrates. Several important comments have to be made with respect to the carbohydrate-binding specificity of lectins. First, plant lectins display a broad range of specificities. Secondly, most lectins have a much higher affinity for oligosaccharides than for simple sugars. Thirdly, structurally different lectins may recognize the same sugars. Fourthly, the specificity of most lectins is directed against foreign (i.e. non-plant origin) glycans.

Plant lectins can be subdivided into so-called specificity groups according to their preferential binding to simple sugars (*Table 2*). Until about a decade ago only mannose/glucose-, Gal/GalNAc-, GlcNAc/$(GlcNAc)_n$-, fucose- and sialic acid-binding lectins were recognized (Goldstein and Poretz, 1986). A novel specificity group was added upon the discovery of the monocot mannose-binding lectins (which exhibit

Table 2. Carbohydrate-binding specificity of plant lectins

Specificity	Example
Fucose-group	
Fucose	*Ulex europaeus* agglutinin I
Galactose/N-acetylgalactosamine-group	
Galactose>>GalNAc	Jacalin (*Artocarpus integrifolia* lectin)
Gal=GalNAc	*Clerodendron trichotomum* agglutinin
Gal<<GalNAc	Soybean (*Glycine max*) agglutinin
N-acetylglucosamine-group	
GlcNAc	Wheatgerm (*Triticum aestivum*) agglutinin
$(GlcNAc)_n$	*Urtica dioica* agglutinin
Mannose-group	
Mannose only	*Galanthus nivalis* agglutinin
Mannose/glucose	ConA (*Canavalia ensiformis* agglutinin)
Mannose/maltose	*Calystegia sepium* agglutinin
Sialic acid-group	
Sialic acid	Wheatgerm (*Triticum aestivum*) agglutinin
Neu5Acα(2,6)Gal/GalNAc	*Sambucus nigra* agglutinin I
Neu5Acα(2,3)Gal/GalNAc	*Maackia amurensis* agglutinin
Complex glycan-group	
With known complex specificity	PHA (*Phaseolus vulgaris* agglutinin)
With unknown complex specificity	*Euonymus europaeus* agglutinin

an exclusive specificity towards mannose) (Shibuya *et al.*, 1988b; Van Damme *et al.*, 1995). In addition, the existence of yet another specificity group was revealed by the recent discovery of a mannose/maltose binding lectin from *Calystegia sepium* (Peumans *et al.*, 1997). Many lectins do not fall within one of these specificity groups because they do not bind any mono- or disaccharide but require more complex glycans for an efficient inhibition of their agglutination activity. They are usually classified as lectins with a 'complex' specificity without any further specification.

The nominal specificity of lectins is usually determined by hapten inhibition assays of the agglutination of cells or the precipitation of glycoproteins. Although a great deal of all known plant lectins can be inhibited by mono- or disaccharides the concentrations required for inhibition are relatively high, especially when compared to the inhibitory concentrations of more complex oligosaccharides. For instance, wheatgerm agglutinin is about 4500-fold better inhibited by $(GlcNAc)_4$ than by GlcNAc (Goldstein and Portez, 1986). Similarly, the snowdrop lectin is 12-fold better inhibited by Manα1,3Man than by Man (Shibuya *et al.*, 1988b). The higher inhibitory potency of the oligosaccharides is due to the fact that the carbohydrate-binding site of the lectins is most complementary to oligosaccharides (and hence has a much higher affinity for the complexer glycans).

The carbohydrate-binding specificity of plant lectins is determined primarily by the three-dimensional structure of the binding site(s). X-ray crystallography, NMR and computer modelling studies have demonstrated that the carbohydrate-binding sites are conserved (in terms of amino acid sequences) within a given lectin family (which does

not necessarily imply that they all have the same specificity) (Rougé *et al.*, 1991; Barre *et al.*, 1996). In addition, these structural analyses demonstrated that the binding sites of the different lectin families have a unique overall structure. This, does not preclude, however, that members from different lectin families can recognize the same sugar. Several carbohydrates are recognized, indeed, by structurally different lectins. For example, mannose-binding lectins occur within the legume lectins, the monocot mannose-binding lectins and the jacalin family. Gal/GalNAc-binding lectins are found in the legume lectin family, the group of type 2 RIP and the amaranthin family. Similarly, $(GlcNAc)_n$-binding lectins occur among the legume lectins, the chitin-binding lectins as well as the Cucurbitaceae phloem lectins. These examples clearly demonstrate that plants developed structurally different motifs to recognize and bind the same mono- or oligosaccharide.

Since one can reasonably assume that the role of lectins is determined by their carbohydrate-binding activity it is important to determine which glycans are recognized best by the carbohydrate binding site of the lectin. Considering that mono- or disaccharides are only weakly bound, and that the binding sites of most lectins can accommodate oligosaccharides, the natural ligands for these proteins are most probably complex glycans rather than simple sugars. This assumption is further supported by the observation that many lectins are exclusively inhibited by complex oligosaccharide side chains of animal glycoproteins. A closer examination of the specificity of plant lectins further indicates that they exhibit a pronounced preference for foreign glycans. Many lectins have an exclusive specificity, indeed, for carbohydrates which are absent in plants. For example, all chitin-binding lectins recognize a polysaccharide that does not occur in plants but is common in fungi as well as in insects. Similarly, the sialic acid-specific lectins bind to a sugar which is not found in plants but is abundant in animal glycoconjugates. The same reasoning holds true for all those lectins with a pronounced preference for complex GalNAc-containing glycans of animal glycoproteins.

THE PHYSIOLOGICAL ROLE OF PLANT LECTINS

In spite of all efforts to elucidate the physiological role(s) of plant lectins, this issue remains controversial. Considering the differences in structure and sugar binding specificity between the lectin families it is rather unlikely that all lectins have the same or a similar function. Most probably, the sugar binding activity and specificity are the determining factors in the normal function of lectins. Moreover, one can expect that specific binding of a lectin to a glycoconjugate receptor (irrespective of whether this receptor is located within or outside the plant) is an essential step in the action of lectins. In the case of plant lectins, receptors are glycoconjugates possessing a carbohydrate moiety with a structure complementary to that of the binding site of the lectin. Evidently, different types of glycoconjugates (glycoproteins, glycolipids and polysaccharides) with identical (or structurally similar) carbohydrates can act as receptors for the same lectin. In order to understand the role of plant lectins, it is important to identify their (potential) receptors. Attempts to identify endogenous receptors (so-called lectin-binders) did not yield significant data. Some lectins have a high affinity for glycoproteins present in the same tissue but it is still questionable whether the observed *in vitro* binding has any physiological relevance. Because of the

apparent lack of endogenous receptors, the search was progressively shifted towards exogenous lectin receptors. Two basic ideas stimulated the search for such exogenous receptors. First, the results of specificity studies pointed towards foreign glycans as the most likely targets of many plant lectins. Secondly, it became increasingly apparent that lectins are the only plant proteins capable of recognizing and binding glycoconjugates present on the surface of microorganisms (i.e. bacteria and fungi) or exposed along the intestinal tract of insect or mammalian herbivores. The broad spectrum of carbohydrate-binding specificities can be regarded as an evolutionary adaptation of the plants to cope with the problems inherent to the recognition of a wide diversity of microbial and animal glycans.

The admission that the activity of most plant lectins is directed against foreign glycans was an important breakthrough in the search of the role of these proteins. Further elucidation of the exact role necessitated studies of the (biological) effects provoked by the lectins upon binding to the receptors in or on their target organism(s). Plant lectins have little, if any, direct effect on viruses, bacteria or fungi. In contrast, several lectins affect the growth and development of insects upon oral uptake. Although the exact working mechanisms are not known, it seems likely that lectin binding to glycan receptors on the surface of the gut provokes some deleterious effects. The oral toxicity of lectins towards higher animals is reasonably well understood as a result of intensive studies of the possible health risk of lectins present in food and feed plants. Most of the present knowledge about the toxic properties of plant lectins and the effects they provoke in animals and men have been obtained from accidental poisoning of humans by raw or insufficiently cooked beans and from animal feeding experiments with purified *Phaseolus vulgaris* agglutinin (PHA) (Pusztai and Palmer, 1977). Upon ingestion, PHA (which is not degraded or inactivated by the stomach or gut digestive enzymes) binds to the brush border cells of the intestine, is rapidly endocytosed and induces an enhanced metabolic activity which eventually leads to hyperplasia and hypertrophy of the small intestine (Pusztai and Bardocz, 1996). In addition, ingestion of PHA or raw beans causes acute nausea followed by vomiting and diarrhoea. Due to the severe discomfort experimental animals are very reluctant to eat a PHA containing diet. The latter observation also demonstrates the repellant activity of PHA, which is essential in view of the physiological role of the lectin. Many plant lectins bind to the intestinal mucosa of rats and disturb the function of the intestine. Moreover, they cause systemic effects such as an enlargement of the pancreas (Pusztai and Bardocz, 1996), which in the long term may be as threatening as the acute toxic effects.

Both the specificity of plant lectins and their deleterious effects on insects and/or higher animals indicate that they play a protective role against predators. Several examples of lectin-mediated resistance are found in nature. For instance, black locust (*Robinia pseudoacacia*) and elderberry (*Sambucus nigra*) accumulate large quantities of bark lectins which cause the same severe toxicity symptoms in rats as PHA. Due to the presence of these lectins, the bark of elderberry and black locust is not attacked by rodents or other animals whereas, in the same habitat, the bark of lectin-free species like poplar, willow and wild apple is very susceptible to the same animals.

Many lectins and especially those found in seeds and vegetative storage tissues closely resemble typical plant storage proteins with respect to their abundance, biochemical properties and developmental regulation. These lectins can be regarded,

indeed, as storage proteins with a particular (i.e. carbohydrate-binding) activity. In this respect, lectins resemble other storage proteins with a well-defined biological activity such as, e.g. protease-inhibitors, α-amylase-inhibitors, toxins. It is generally believed, therefore, that the storage-protein-like lectins have a dual role. As long as the plant is not challenged the lectins do not fulfil any specific role but just represent a store of nitrogen. When the plant (or a tissue) is attacked by insects or higher animals, the lectins act as aspecific defence proteins. The incorporation of (storage) nitrogen under the form of toxic or repellant proteins in seeds and vegetative storage organs can be seen as an evolutionary adaptation. Most probably, the presence of toxic lectins offers no complete protection to a seed or plant tissue against consumption. However, the reaction of avoidance by the animal may be beneficial for the survival of the species (Peumans and Van Damme, 1995).

The above described defensive role holds true for many lectins from different lectin families but can certainly not be generalized. Lectins occurring at low concentrations may well be involved in more specific recognition processes either within or outside the plant. For example, legume root lectins may play a role in the recognition and/or binding of *Rhizobium* and *Bradyrhizobium* species (Bohlool and Schmidt, 1974; Diaz *et al.*, 1989).

IMPACT OF THE PRESENCE OF LECTINS IN CROP PLANTS ON THE SAFETY AND QUALITY OF THE FOOD AND FEED PRODUCTS

The obvious toxicity of plant lectins towards predatory invertebrates and higher animals raises the question of the potential risks inherent to the presence of lectins in crop plants destined for the production of food and/or feed products. Modern human nutrition largely depends on a limited number of crop plants irrespective whether they are consumed directly or indirectly (i.e. after conversion into protein rich products by husbandry animals). In the past, lectins have unambiguously been identified in numerous edible plants including cereals, legumes, vegetables and fruits (*Table 3*). To assess the possible risk(s) of the presence of a lectin, several factors have to be taken into account. When the plant materials are eaten raw, the intrinsic toxicity and concentration of the lectin, and the possible presence of antagonistic or synergistic compounds determine the risk(s). In principle, the same holds true for processed food or feed products except that not the total lectin concentration, but the level of biologically active lectin must be taken into account. Since the latter depends on the survival rate of the raw material during the processing step(s), the intrinsic stability of the lectin is an important factor in the assessment of the possible health risks of lectins in processed food products (Peumans and Van Damme, 1996).

For obvious ethic reasons the possible adverse effects of dietary lectins on human beings are only documented by observations of accidental poisonings. Ingestion of a few raw castor beans causes death within a few days because of the presence of ricin, a highly toxic type 2 ribosome-inactivating protein. Besides type 2 RIP, no other plant lectins have direct lethal effects. Severe gastrointestinal distress is caused by the kidney bean (*Phaseolus vulgaris*) lectin PHA upon accidental poisoning by insufficiently cooked beans. Although a short exposure to this legume lectin is not lethal and the effects are reversible, a prolonged exposure may eventually cause death (e.g. as a result of a severe diarrhoea). In the absence of experimental evidence obtained with

Table 3. Food plants with high lectin levels in edible parts

Food plant	Lectin present in	Maximal lectin concentration (g/kg)	Type of lectin	Oral toxicity of the lectin
Legumes				
Broad bean	Seed	<1	Legume lectin	Low
Kidney bean	Seed	<10	Legume lectin	High
Lentil	Seed	<1	Legume lectin	Low
Pea	Seed	<1	Legume lectin	Low
Peanut	Seed	<2	Legume lectin	Low
Soybean	Seed	<2	Legume lectin	Moderate
Cereals				
Barley	Germ	<0.5	Chitin-binding lectin	Not determined
Rice	Germ	<0.5	Chitin-binding lectin	Not determined
Rye	Germ	<0.5	Chitin-binding lectin	Not determined
Wheat	Germ	<0.5	Chitin-binding lectin	Moderate
Grain amaranth	Seed	<1	Amaranthin	Not determined
Vegetables				
Garlic	Bulb	<2	MMBL[a]	Not determined
Leek	Leaf	<0.01	MMBL	Not determined
Onion	Bulb, leaf	<0.01	MMBL	Not determined
Shallot	Bulb, leaf	<0.01	MMBL	Not determined
Taro	Tuber	<5	MMBL	Not determined
Potato	Tuber	<0.05	Chitin-binding lectin	Not determined
Tomato	Fruit	<0.01	Chitin-binding lectin	Not determined
Cucumber	Fruit	<0.01	C.phloem lectin[b]	Not determined
Melon	Fruit	<0.01	C.phloem lectin	Not determined
Pumpkin	Fruit	<0.01	C.phloem lectin	Not determined
Fruits				
Jackfruit	Seed	<2	Jacalin	Not determined
Elderberry	Fruit	<2	Type 2 RIP	Not determined
Oil plants				
Castor bean	Seed	<5	Type 2 RIP	Lethal

[a] Monocot mannose-binding lectin
[b] Cucurbitaceae phloem lectin

humans, the possible adverse effects of lectins are usually inferred from experiments with laboratory animals like mice, rats or pigs. Although extrapolation from animal models to humans must be done with care, the observed adverse effects of many dietary lectins on the gut and other organs of animals are of great concern in view of food safety. An additional problem is the fact that the toxicity of dietary lectins is usually studied in acute or short-term experiments. As a result, little is known about the effects caused by a constant or intermittent exposure to a dietary lectin over a long period. Taking into consideration that most lectins have mitogenic activity towards different cell types, induction of benign or malignant tumors by dietary lectins may be possible. The fact that some lectins end up in the circulatory system is even more worrying since a lectin-mediated induction of uncontrolled cell growth can take place in any organ.

Lectins are common in crop plants (Jaffé, 1980; Nachbar and Oppenheim, 1980; Liener, 1986) (*Table 3*). Most legume food (and feed) plants contain lectins in their

seeds. In general legume lectins are heat-labile and do not survive processing. Some *Phaseolus* species, however, contain high levels of toxic lectins with a fairly high heat stability. Unless properly processed, these beans may contain some residual lectin activity. Several important staple crops like cereals and potatoes, as well as vegetables like tomatoes contain heat-stable chitin-binding lectins which can survive processing. Since it has been demonstrated that, for instance, wheatgerm agglutinin definitely exhibits harmful effects in rats (Pusztai *et al.*, 1993a), the presence of chitin-binding lectins in food plants (especially when they are eaten raw) is of some concern. Type 2 ribosome-inactivating proteins are not very common in crop plants. Ricin occurs in large quantities in castor beans but properly processed castor bean oil is essentially free of the toxin. Elderberry fruits, which in some countries are used for the production of juices, also contain type 2 RIP. Although these RIP are far less toxic than ricin they are fairly heat stable and therefore deserve some attention. Several food plants (e.g. garlic, ramsons) contain high levels of monocot mannose-binding lectins in their edible parts. Since many of these species are eaten raw, biologically active *Allium* lectins are common constituents of the human diet. Although there is no evidence for toxic effects of monocot mannose-binding lectins, their low mitogenic activity may be of some concern. It should be mentioned, however, that the dietary monocot mannose-binding lectins could be beneficial since they reduce bacterial growth in the gut (Pusztai *et al.*, 1993b). At present, no data are available about the possible adverse effect of several other important dietary lectins like amaranthin (from *Amaranthus caudatus* seeds) and the lectins from bananas and pumpkins. No toxic effects have been described, but the presence of lectins especially in the latter two species may be of some concern (as they are eaten raw).

Some attention should also be given to the occurrence of yet unknown dietary lectins. One cannot exclude that some lectins present in crop plants escaped detection. In addition, the presence of low levels of a toxic lectin can be masked by the occurrence of high concentrations of a harmless lectin. Similarly, some of the recently introduced novel vegetables and fruits may contain toxic lectins.

Crop plants contain besides lectins also other types of defence proteins like enzyme-inhibitors, lytic enzymes, type 1 RIP, thionins and pathogenesis-related proteins. Since these proteins act through different working mechanisms, there is a possibility that the adverse effects of dietary lectins are reinforced by an otherwise harmless protein or vice versa. The observation that a combination of the lectin and protease inhibitor from legume seeds (e.g. soybean) exhibits a stronger antinutritional effect in laboratory animals than the individual proteins certainly points in this direction.

Applications in biotechnology

At present, lectins are undoubtedly the most versatile group of plant proteins used in basic and applied biological and biomedical research. Although the overwhelming success of plant lectins is based in the first place on their highly specific carbohydrate-binding activity and the biological effects they provoke in various organisms, the availability of reasonable quantities of pure preparations of many lectins also contributed to their success. In the past the exploitation of plant lectins was restricted to native and/or chemically modified natural products (i.e. lectins isolated from plant materials), and therefore was subject to considerable limitations. During the last decade,

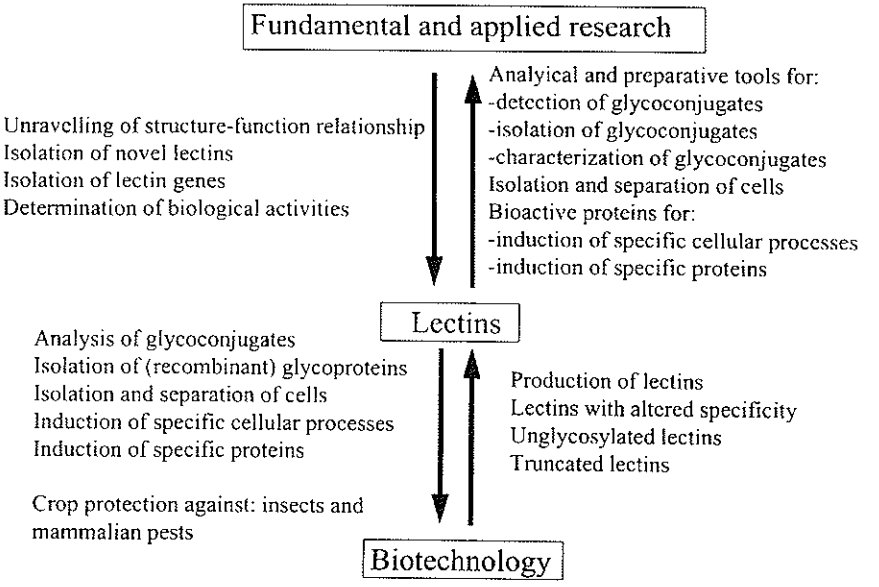

Figure 1. Interplay between lectinology, research and biotechnology.

important progress has been made in the molecular cloning and analysis of plant lectin genes as well as the expression of lectins in heterologous systems. These new developments offer interesting perspectives not only for a more intensive use of lectins but also for novel applications in other fields (*Figure 1*).

The broad range of potential applications of lectins in biotechnology are situated in three different arbitrarily delineated areas. First, lectins will be used as tools – in a broad sense – in biotechnological processes. Secondly, recombinant DNA technology and protein engineering will improve the availability and/or properties of lectins. Thirdly, lectins offer interesting perspectives to improve the resistance or quality of transgenic (crop) plants (*Table 4*).

APPLICATIONS OF LECTINS AS TOOLS AND AS BIOACTIVE PROTEINS IN BIOTECHNOLOGY

Purified plant lectins are used in biotechnology as tools and as bioactive proteins. In the former case the applications are based on the specific carbohydrate-binding activity of the lectins whereas in the latter case the biological effects of lectins on cells, tissues or whole organisms are exploited. Applications as tools are reasonably well understood in molecular terms because they rely on a well-defined interaction between a lectin and its carbohydrate ligand. Once the target carbohydrate structures are known, experimental procedures can easily be designed with a predictable outcome. The exploitation of the biological effects caused by lectins also relies on a specific interaction between the lectins and carbohydrate-receptors (on the cell membranes) but involves additional signaling reactions within the target cells. Since the sequence of events between the binding of the lectin and the appearance of the

Table 4. Important applications of plant lectins

I.	In research
A.	As tools
	Detection of specific glycans
	Isolation of glycoconjugates
	Characterization of glycans
B.	As bioactive proteins
	Induction of mitosis
	Induction of specific proteins or cellular processes
II.	Biomedical applications
A.	Diagnostic
	Blood group typing
	Tracing of aberrant glycosylation of glycoproteins
	Histochemical staining of carbohydrates
	Stimulation of lymphocytes for chromosome analysis
B.	Therapeutic
	Immunomodulation
	Cancer therapy with immunotoxins
III.	In plant protection
	As resistance factors against predators

biological effect(s) are not fully understood in molecular terms, the use of lectins as bioactive proteins is still in an experimental stage.

Applications of lectins as bioactive proteins

Plant lectins are intensively used as inducers of specific processes in animal or human cells. Well-known applications are the activation of lymphocytes with the so-called mitogenic lectins and the induction of the synthesis of specific proteins like enzymes, cytokines or interleukines (Kilpatrick, 1991). At present, only a few lectins like ConA and PHA are fully exploited to elicit specific cellular reactions. Many other plant lectins which also have interesting biological activities, and perhaps are superior to the classical lectins for various applications are not used or only on a very limited scale. The exploitation of the unique biological activities of lectins can certainly be improved by a more intensive application of the currently available 'lectin technology'. In addition, further corroboration of the biological effects of old and novel lectins on different cell types eventually will lead to a whole range of yet unknown applications. A few examples illustrate the underexploitation of the currently available lectins. The lectins from tulip bulbs and *Colchicum autumnale* (meadow saffron) tubers, for instance, have interesting mitogenic properties (Bemer *et al.*, 1996) but are not commonly used. Similarly, the superantigenic properties of the lectin from stinging nettle (*Urtica dioica*) are not yet exploited (Galleli and Truffa-Bachi, 1993). The importance of further research on lectins is illustrated by the recent discovery of a novel lectin from *Calystegia sepium* (hedge bindweed). This lectin is a potent mitogen but is less cytotoxic than ConA or PHA and hence can be used in a much broader range of concentrations (Peumans *et al.*, 1997). Since all four of the above mentioned lectins are either commercially available or can easily be isolated, the exploitation of their unique properties can be achieved on a short term.

Some lectins not only affect isolated cells but also provoke specific effects on whole organisms upon injection or oral ingestion. In general, the observed effects are the result of a (mitotic) stimulation of cells in the gut or in another organ. As discussed above, the administration of lectins usually has a negative effect on the health of the organism. However, in some cases lectins may exert beneficial effects. For example, some lectins stimulate the immune system through their mitogenic activity on lymphocytes. In this respect intensive efforts are undertaken to corroborate the application of MLI (one of the lectins from *Viscum album*) as an immune stimulant in cancer therapy (Gabius *et al.*, 1992). At present, it is not clear whether lectins can induce the (over)production of valuable biomolecules in higher organisms. If so, laboratory animals can be used to produce interesting macromolecules without the need for introduction of foreign genes.

Applications as tools

Detection and isolation of glycoconjugates. In principle, lectins can be used to detect and/or isolate any glycoconjugate (e.g. glycoproteins, polysaccharides, glycolipids) containing a carbohydrate with a structure that is complementary to the binding site(s) of a lectin (Cummings, 1997). Due to the broad range of specificities covered by plant lectins, powerful lectin-based analytical and preparative detection and isolation techniques are now available for most glycans. Carbohydrate detection techniques are usually based on the binding of a purified lectin on specific carbohydrates present on cells, tissue sections or (isolated) glycoconjugates, followed by a visualization of the bound lectin. Visualization can be direct (e.g. when fluorescent-labelled lectins are used) or indirect through the application of immunological or other techniques (e.g. using the biotin-streptavidin method). To check the specificity of the lectin staining control experiments with an excess of free ligand are routinely included. In addition to basic and applied research, carbohydrate detection techniques are intensively used in histochemistry and histopathology (Schumacher *et al.*, 1991; Gabius and Gabius, 1991). Applications in histopathology are usually based on the expression of aberrant glycans on the surface of transformed cells. The detection of such aberrant glycans by lectins is often indicative for an (early) diagnosis. Aberrant glycans are not only expressed on cell surfaces but can also occur on serum proteins (e.g. as a result of a defective glycosyltransferase). Once the relevant protein has been purified, the possible occurrence of aberrant glycans can be checked with a panel of lectins. The detection of specific carbohydrates on glycoconjugates is also important in biotechnology. Many bioactive proteins require a correct glycosylation for *in vivo* activity and/or stability. As a result, recombinant glycoproteins produced by animal or human cells can only be used if their glycan chains have the correct structure. Lectins can help to check the accuracy of glycosylation. For example, the occurrence of a terminal Neu5Acα(2,6)Gal/GalNAc (which is often required for activity or stability) can easily be checked with the elderberry (*Sambucus nigra*) lectin SNAI (Shibuya *et al.*, 1987) (*Figure 2*).

Plant lectins are also powerful tools for the isolation and fractionation of glycoconjugates and for the study of oligosaccharides and glycopeptides (Osawa and Tsuji, 1987). Virtually all modern purification schemes exploit the specific and

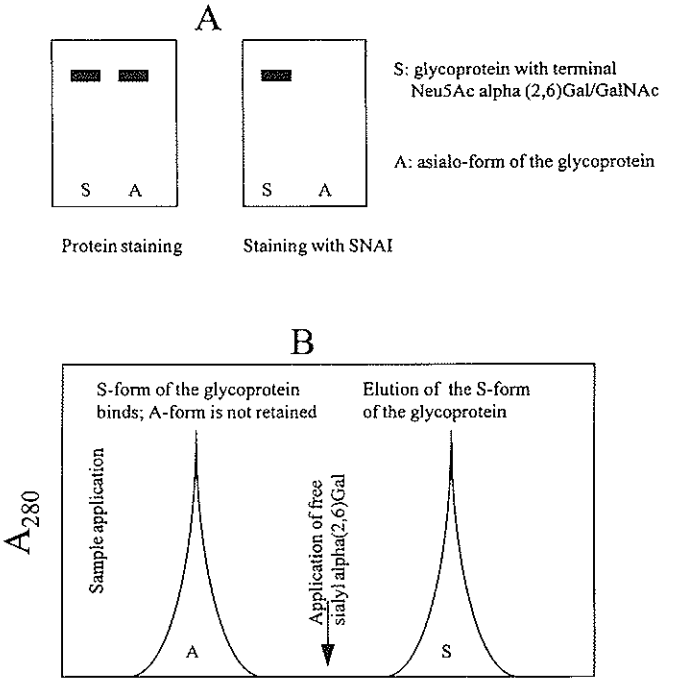

Figure 2. Use of *Sambucus nigra* agglutinin I (SNAI) for (A) the detection of terminal Neu5Acα(2,6)Gal/GalNAc residues on glycoproteins and (B) the separation of sialylated and unsialylated forms of the same glycoprotein.

reversible binding of glyconjugates to a lectin immobilized on an inert matrix (Green and Baenzinger, 1989; Kobata and Endo, 1992). In principle, lectin-affinity chromatography is simple and straightforward. A crude or partially purified culture medium, cell lysate or tissue homogenate is applied on to a column filled with an appropriate matrix containing the immobilized lectin. All glycoconjugates carrying a glycan which is recognized by the lectin are retained on the column. After washing the column with buffer, the bound glycoconjugates are displaced with an excess of free ligand (usually a mono- or oligosaccharide). Thereafter, the affinity-purified compounds can be recovered by conventional techniques and subjected to further separation. Lectin-affinity chromatography is often associated with immobilized ConA because this particular lectin binds most, but not all, glycoproteins. In principle, chromatography on ConA allows only a separation between glycosylated and unglycosylated proteins. Many other plant lectins have a narrower specificity than ConA and bind only a fraction of the glycoproteins. The selectivity of lectin affinity chromatography is illustrated by the binding of human serum proteins to the mannose-specific lectins ConA and GNA. Immobilized ConA binds most human serum glycoproteins whereas only α-2 macroglobulin is retained on immobilized GNA (Shibuya *et al.*, 1988a). This example not only demonstrates that the nominal specificity of a lectin is of little predictive value but also underlines the extremely high selectivity one can obtain when an appropriate lectin is available.

Lectin-affinity chromatography is a very powerful and versatile technique for the

purification of glycoconjugates, and offers many advantages. First, the availability of many lectins with different specificities makes this technique applicable for virtually any glycan. Secondly, a consecutive use of two or more lectins with different specificities allows a rapid separation of glycoconjugates with different glycan chains Thirdly, lectin-affinity chromatography can be introduced at any stage of the purification of a glycoconjugate and requires no special pretreatment of the fractions to be purified. Since the lectin–glycoconjugate interaction is, at least within certain limits, insensitive to the presence of salts (including ammonium sulphate) only the pH of an extract or a partially purified fraction has to be adjusted. Fourthly, lectin-affinity chromatography usually yields a high purification factor, especially when lectins with a highly selective specificity towards the target glycoconjugates can be used. Fifthly, lectin affinity chromatography can be used to concentrate glycoconjugates from dilute solutions (e.g. from large volumes of a culture medium). Sixthly, the overall yield of lectin-based affinity chromatography is high (at least when performed under optimal conditions). Seventhly, affinity-purified fractions desorbed from a column of immobilized lectins can be further processed by conventional protein purification techniques without complex or time-consuming pretreatment. Eighth, lectin-affinity chromatography can easily be scaled up or scaled down.

Lectin affinity chromatography is undoubtedly of great potential value in biotechnology since it offers interesting perspectives for the purification of (recombinant) glycoproteins produced by eukaryotic cells. The high selectivity coupled to a high recovery, and the possibility of processing large volumes without time-consuming or expensive pretreatment are important advantages for industrial or semi-industrial processes. At present, there are still some practical problems which may hamper a large scale use of lectin-affinity chromatography. The limited availability and/or high prices of suitable lectins makes the preparation of large quantities of immobilized lectins expensive. In some instances, the coupling of lectins to inert matrices has not yet been optimized. Regeneration of the columns without affecting the activity of the immobilized lectins can be difficult or even make a repeated use impossible. Most of these practical problems will probably be solved in the near future because the basic knowledge of the structure and function of plant lectins will lead to an improved exploitation of their unique properties. In addition, the limited availability of some lectins can be solved by a (plant) biotechnological approach.

IMPROVING THE AVAILABILITY AND PROPERTIES OF PLANT LECTINS THROUGH BIOTECHNOLOGY

Increasing the availability of lectins with interesting properties through recombinant DNA technology

Commercial availability of a protein at a reasonable price is a prerequisite for its intensive use as a research tool. Although the preparation of plant lectins is, in principle, relatively easy because powerful affinity chromatography techniques can be included in the purification schemes, many interesting lectins can not be prepared on a large enough scale for practical reasons. In some cases the starting material is not available in sufficient quantities. For instance, the preparation of large quantities of the

orchid lectins (which exhibit a very interesting specificity and have unique biological activities) is virtually excluded because these species are protected. In other cases, the lectin concentration in the plant materials is so low that the costs for the preparation of the purified proteins impedes their commercialization. The latter holds true, for instance, for the lectins from onions or leek. Even when the lectin concentration in a given plant tissue is reasonably high, the preparation of the pure protein may be very cumbersome because of the presence of large quantities of polysaccharides, phenolic compounds or other interfering substances. The complex and costly procedures required for the purification of these lectins makes them unattractive for large-scale isolation. For instance, problems caused by high levels of pectic substances and anthocyanins in elderberry fruits make a cost-effective purification of the fruit-specific lectin virtually impossible.

Special problems arise when a plant contains two or more (iso)lectins, which even when present in sufficient amounts, can not be prepared in large quantities because the currently available techniques do not allow a preparative separation of the different (iso)lectins. When the different isoforms of a given lectin have the same specificity and exhibit the same biological activities, a total lectin preparation can be used without any further purification. However, when there are differences in specificity and/or biological activities between the isolectins, there is an obvious need for highly purified isoforms. In some cases, individual isolectins can readily be separated from a total lectin preparation. For instance, the leucoagglutinating (L4) and erythroagglutinating (E4) isoforms of PHA can easily be obtained by ion exchange chromatography (Feldsted *et al.*, 1977). Similarly, the leukoagglutinating (MAL) and haemagglutinating (MAH) forms of the *Maackia amurensis* seed lectin can be separated from each other by ion exchange chromatography. In other cases, however, it is virtually impossible to isolate single isolectins. For example, most, if not all, monocot mannose-binding lectins are encoded by extended gene families. Detailed biochemical and molecular analyses have demonstrated, indeed, that the lectins from daffodil and snowdrop probably comprise more than 100 isoforms. Although all these isoforms have a similar overall specificity and share a high sequence similarity, it is questionable whether they all exhibit identical biological activities. The latter question is not only academic but is also of great practical importance in view of the envisaged use of these lectins as insect resistance factors in transgenic crop plants. To identify the most potent entomotoxic isolectin, individual isoforms should be tested. Unfortunately, this test cannot be performed because single isolectins can not be prepared yet.

In all these cases where the isolation of a potentially useful lectin cannot be performed because of practical reasons, a biotechnological approach can solve the problem. The cloned lectin genes can be transferred to, and expressed in, a heterologous system. Transgenic plants, yeast as well as bacteria can be used as an expression system for plant lectins. Bacterial expression systems offer several important advantages. Their transformation is relatively easy and cultures can be grown and maintained under controlled laboratory conditions. However, the recombinant proteins are rather difficult to isolate from the bacteria since they usually are deposited in an insoluble form in protein inclusion bodies. In addition, it is not evident that the bacterial recombinant lectins have the same properties as the original plant lectins. Most mature plant lectins are the result of a (sometimes complex) set of post-translational modifications. Bacteria are unable to perform some of these modifications (e.g. glycosylation,

proteolytic cleavage, intermolecular disulphide bridge formation) and, therefore, may accumulate a recombinant lectin that differs from the native plant lectin. Transgenic plants are probably better candidates for a large-scale production of recombinant lectins. Although the transformation technology of plants and the growth and maintenance of the transgenics is certainly more complicated than that of bacteria, plant systems offer several important advantages. First, less problems will be encountered with the post-translational processing and modifications of the recombinant lectins. Secondly, apart from regulatory restrictions, transgenic plants can be grown on a large scale at low cost. Thirdly, the recombinant lectins are present in a soluble (i.e. water- or buffer extractable) form and therefore can be extracted by gentle procedures. The efficiency of the production of recombinant lectins in transgenic plants will be determined primarily by the choice of the expression system. Ideally, the lectins should be expressed in an otherwise lectin-free tissue with a high protein level and low contents of substance(s) which interfere with the isolation of the recombinant lectins. Taking into consideration the present state of the art, seeds of oil seed rape (canola) are probably the best suited expression system. The transformation technology of this plant is well established and high expression levels of recombinant proteins can be obtained in the seeds through the use of its own 2S seed albumin promoter. Canola seeds are rich in oil and protein but contain no endogenous lectin. Aqueous extracts from defatted (by cold pressing or solvent extraction) seed meal have a high protein content and are easy to handle. Due to the absence of excess polysaccharides or other interfering compounds, the recombinant lectins can be purified following the procedures developed for the isolation of the native lectins. A canola crop yields about 5 tonnes of seeds per ha with an overall protein content of about 20%. Assuming an expression level of 1% of the total seed protein, the potential yield of transgenic lectins amounts to about 10 kg/ha.

Improving the properties of lectins through protein engineering

NMR and X-ray diffraction analyses enabled to determine the three-dimensional structure of the carbohydrate-binding binding sites of several families of plant lectins. In addition, molecular modelling allowed to make fairly accurate predictions of the 3D structure of the binding site(s) of many other lectins based on their (deduced) amino acid sequences and the atomic coordinates of the model lectins. Binding of a carbohydrate to a lectin is mediated by a complex network of hydrogen bonds and hydrophobic interactions between the side chains of the amino acid residues comprising the binding site and the sugar residue(s) (Weis and Drickamer, 1996). A perfect match between the structure of the carbohydrate and the binding site of the lectin is a prerequisite for binding. Changes in the 3D structure of the binding site (e.g. by substitution of one of the amino acid residues involved in the binding site) usually have dramatic negative effects on the activity of the lectin. This explains why, at least within a lectin family, the amino acid residues composing the sugar binding site(s) are highly conserved. In those cases where an amino acid substitution has taken place in the binding site the resulting protein may be completely devoid of carbohydrate-binding activity. Examples of so-called lectin-related proteins have already been found in the legume lectins (Mirkov *et al.*, 1994; Van Damme *et al.*, 1995a), the monocot mannose-binding lectins (Barre *et al.*, 1996) and the type 2 RIP (Van

Damme *et al.*, 1997a). Amino acid substitutions in the binding site do not necessarily completely abolish the carbohydrate-binding activity of a lectin. For instance, when a lectin polypeptide harbours two or more binding sites, inactivation of a single site usually leaves the other site(s) unaffected. The resulting lectin may have a reduced activity but still recognizes and binds a specific carbohydrate. Molecular modelling of the monocot mannose-binding lectins, for example, demonstrated that the number of active sites per monomer varies between 0 and 3, and confirmed that there is a good correlation between the number of active binding sites and the biological activity (Barre *et al.*, 1996). In addition, the same modelling studies indicated that the obvious differences in biological activity and fine specificity between the lectins with three active sites per monomer can be attributed to subtle changes in the 3D structure of the binding sites. Lectins with a single binding site per polypeptide chain are, in principle, more sensitive to amino acid substitutions in the carbohydrate binding site itself or in the vicinity of this site. For example, the insertion of three extra amino acid residues in the binding site of the lectin from *Cladrastis lutea* yields a completely inactive lectin-related protein (Van Damme *et al.*, 1995a). Site-directed mutagenesis of several legume lectins confirmed that alterations in the binding site often result in inactive or less active lectins (van Eijsden *et al.*, 1992). However, diversity among naturally occurring lectins demonstrates that amino acid substitutions in the lectin polypeptides can also have positive effects. For instance, the wide specificity range found within the legume lectin family is based on subtle differences in the 3D structure of small parts (i.e. the binding sites) of otherwise very similar proteins. Similarly, the differences in fine specificity between the type 2 RIP also rely on small differences in the 3D structure of the binding sites. The latter examples demonstrate that even subtle differences in the 3D structure of the binding site of a lectin can have profound effects on its activity and/or specificity. Given the rapidly expanding knowledge of the 3D structure of lectins and the increasing power of molecular modelling, protein engineering can be used to design and produce lectins with improved properties (i.e. better suited for practical purposes in research and biotechnology).

The potential of lectin engineering can be illustrated by a few practical examples. Legume lectins, which show a wide natural variation in specificity, are probably the best candidates for the development of lectins with altered or even novel specificities. Using the information available on the 3D structure of legume lectins, one can try to design binding sites complementary to for example xylose. Since xylose-specific lectins are still unknown, such a lectin would be very useful. Similarly, it may be possible to develop lectins with an absolute specificity towards well-defined di-, tri- or tetrasaccharides (or even more complex oligosaccharides). The availability of such lectins certainly will improve the use of lectins as tools. A second example of lectin engineering concerns the type 2 RIP. The B chain of a type 2 RIP possesses two dissimilar binding sites. Since both sites have a preference for different sugars the specificity of type 2 RIP is not really clear-cut. Construction, through protein engineering, of type 2 RIP with two identical sites (or alternatively with a single active site) will yield lectins with a homogeneous specificity. This approach can be followed to develop an improved Neu5Acα(2,6)Gal/GalNAc-specific lectin starting from the elderberry (*Sambucus nigra*) lectin SNAI. The B chain of the latter type 2 RIP (which is the only known Neu5Acα(2,6)Gal/GalNAc-specific plant lectin) harbours a Neu5Acα(2,6)Gal/GalNAc-binding site (domain 2) and a Gal-binding site (domain 1)

(Van Damme *et al.*, 1996b). Replacing the binding site of domain 1 by a site identical to that of domain 2 will give a lectin with an exclusive specificity towards Neu5Acα(2,6)Gal/GalNAc. Due to its 'homogenous' specificity, the engineered SNAI will be superior to the original SNAI for all applications based on a selective recognition and binding of Neu5Acα(2,6)Gal/GalNAc. SNAI is only one example of a type 2 RIP that can be improved by engineering of the carbohydrate-binding sites. A similar reasoning holds true for type 2 RIP used for the construction of immunotoxins. Homogenizing the specificity of these RIP will certainly increase the selectivity of the immunotoxins and improve their therapeutic value (Pastan *et al.*, 1992).

Special applications

Production of unglycosylated lectins. Many plant lectins are glycoproteins carrying one or more asparagine-linked glycan chains. In addition, there are also a few examples of plant lectins (e.g. the Solanaceae lectins) with serine and/or hydroxyproline-linked O-glycan chains (Allen *et al.*, 1996). Most probably the carbohydrate moiety of the glycosylated lectin does not directly affect their activity. Indirect effects, however, are likely because the glycan chains change the solubility and perhaps also the stability of the lectin. Since (naturally) glycosylated lectins are perfectly suited as tools, there is no need for unglycosylated variants for most applications. However, when a lectin is injected in animals or humans (e.g. in therapies based on the use of lectins or lectin-containing compounds such as immunotoxins) glycosylation becomes an important issue because plant glycans are potent antigens. Due to the strong response of the immune system, glycosylated lectins can not be injected repeatedly, and therefore, must be replaced by the unglycosylated forms. Since chemical or enzymatic deglycosylation of native lectins is difficult and impractical, unglycosylated variants of lectins can only be obtained in reasonable quantities through protein engineering and recombinant DNA technology. In principle there are two alternative solutions. First, the unmodified (at least in the coding part) lectin gene is expressed in bacteria (which lack the glycosylation machinery). Secondly, the potential glycosylation sites are knocked out by site directed mutagenesis of the lectin gene and the modified gene is expressed in a heterologous eukaryotic system.

Production of truncated lectins. Most plant lectins require an intact (or slightly modified) polypeptide for activity because their carbohydrate-binding site(s) comprise(s) amino acid residues dispersed over the protomer. Chimerolectins, however, contain a domain which is not involved in sugar binding and can be removed without any effect on the carbohydrate-binding activity. Several lectins are processing products of chimeric precursor molecules. The chitin-binding lectins hevein and UDA, the lectin from stinging nettle (*Urtica dioica*) are the (N-terminal) remnants of large chitinase-like precursors (Lerner and Raikhel, 1992; Lee *et al.*, 1991). Similarly, SNAII, a GalNAc-specific lectin from elderberry (*Sambucus nigra*) is derived from a precursor of a type 2 RIP (Van Damme *et al.*, 1996c). It should be mentioned that in the latter case a single precursor is converted into a typical type 2 RIP (namely SNAV or nigrin b) and a lectin composed of two B chains. SNAII has the same specificity and agglutination properties as SNAV but has no *in vivo* or *in vitro* toxicity (because of the

absence of the A chain), and thus is preferred over SNAV whenever the N-glycosidase activity is unwanted. In principle, type 2 RIP can be separated into the A and B chain (after reduction of the intra-chain disulphide bridge). However, this separation is cumbersome and a complete resolution can hardly be achieved. Recombinant DNA technology offers an alternative approach to obtain pure B chains (or A chains). The A domain can be excised from the original type 2 RIP gene and the resulting truncated construct expressed in bacteria or transgenic plants. This technique has already successfully been applied for the synthesis of the A chain of a type 2 RIP from mistletoe (*Viscum album*).

EXPLOITATION OF THE DEFENCE PROPERTIES OF LECTINS IN PLANT PROTECTION

Because of their presumed involvement in plant defence mechanisms intensive efforts are undertaken to assess the use of lectin genes as resistance factors against plant pathogens and predators. As has been outlined above, there is little conclusive evidence for direct effects of lectins on plant viruses or plant pathogenic microorganisms. Some type 2 RIP definitely reduce the number of necrotic lesions upon infection of leaves with viruses but it is still questionable whether the observed protection is due to a direct effect on the virus (e.g. through a deadenylation of the viral RNA) or to a hypersensitive response as a result of RIP-mediated cell death. Lectins are also poor candidates for a direct defensive role against pathogenic bacteria. Apart from indirect effects resulting from the immobilization of bacteria, as has been observed for the seed lectin from thorn apple (*Datura stramonium*) (Broekaert *et al.*, 1986), or a type 2 RIP-mediated hypersensitive response, lectins can hardly affect the growth of (pathogenic) microbes because of the bacterial cell wall barrier. The fungal cell wall also is a barrier for most plant lectins. Even the chitin-binding lectins, which by virtue of their specificity can bind to the major cell wall component of most fungi, can not freely move through the cell wall, and therefore have little if any effect on the growth and development of fungi. There are, however, a few exceptions. For example, class I chitinases are potent antifungal proteins both *in vivo* and *in vitro*, and confer an increased resistance against fungi in transgenic plants (Broglie *et al.*, 1991). Class I chitinases can be considered as chimerolectins (composed of a hevein domain and a catalytic domain) but are better known as group 3 pathogenesis-related proteins. For this reason, a discussion of class I chitinases falls beyond the scope of this contribution. Moreover, this subject has excellently been reviewed on several occasions (Linthorst, 1991; Collinge *et al.*, 1993). Besides the class I chitinases, some other chitin-binding proteins exhibit antifungal activity. UDA, the lectin from rhizomes of stinging nettle (*Urtica dioica*), definitely has fungistatic properties (Broekaert *et al.*, 1989). UDA does not interfere with the metabolism of the fungal cells but disturbs the synthesis and/or deposition of chitin in the cell wall (Van Parijs *et al.*, 1992). UDA is a less potent antifungal protein than the class I chitinases but clearly enhances the effect of the latter enzymes. Therefore, UDA is certainly a valuable candidate to increase the fungal resistance of transgenic plants. Several chitin-binding merolectins composed of a single chitin-binding domain behave as antifungal proteins. Hevein, a latex protein of the rubber tree (*Hevea brasiliensis*) is slightly less active than UDA (Van Parijs *et al.*, 1991). In contrast, the 30 amino acid chitin-binding polypeptide from *Amaranthus caudatus* seeds is a reasonably potent antifungal (but not fungicidal)

agent and may be a valuable candidate to confer or increase fungal resistance in transgenic plants (Broekaert *et al.*, 1992).

Plant eating organisms are the most obvious targets of defence-related lectins because of the abundance of potential glycan receptors on the membranes at the luminal surface along their intestinal tract. Although binding of a lectin to a glycoprotein receptor may take place without further consequences, it is a prerequisite for activity. The reactions that take place after binding to the receptors are determined by the intrinsic properties of the lectin and vary in severity as well as in timing. Some lectins have only mild effects after prolonged exposure whereas others will almost instantaneously cause very severe symptoms (Pusztai and Bardocz, 1996). Although the severity of the symptoms following the ingestion of a lectin is an important factor, the protective power of a lectin depends also on the timing of the appearance of the symptoms. A clear distinction has to made, indeed, between toxic and deterrent/repellant lectins. Toxic lectins cause deleterious effects in the target organisms but the symptoms occur with a delay of several hours or days. Depending on the severity of the effects, feeding continues for a shorter or longer period. For example, ricin intoxication symptoms appear only 12 h after ingestion. As a result, ricin gives no immediate protection against predators since feeding can continue for at least several hours. Deterrent/repellant lectins, on the contrary, provoke severe symptoms almost instantaneously and prevent the animal from further eating. For example, ingestion of PHA or raw beans causes acute nausea followed by vomiting and diarrhoea. Since the severe discomfort withholds the animals from eating, PHA offers an immediate protection for the plant. The concept of toxic and deterrent/repellant lectins is based on experimental evidence obtained with higher animals. It has not been demonstrated yet that this concept applies also to phytophagous invertebrates but it seems likely based on the analogy of the observations made in animal systems.

In principle, lectin genes are good candidates to confer resistance to phytophagous invertebrates and herbivorous higher animals. However, basic differences between the feeding behavior of higher and lower animals and the impact thereof on the host plant necessitate a different approach to develop an efficient protection system against either of the two groups of organisms.

Herbivorous animals can be life-threatening to plants because of their continuous high demands for plant-based food stuffs. Evolutionary adaptations protect most plants, at least in their natural habitat, against excessive damage by herbivores. Besides physical adaptations (e.g. thorns, spikes) the accumulation of toxic or repellant compounds is a common defence mechanism against herbivores. Although most toxins and repellants are low molecular weight compounds (e.g. alkaloids, flavonoids), proteins also act as defence molecules. Lethal plant proteins are rather exceptional. Only a few type 2 RIP like ricin and abrin cause death upon oral ingestion. Moderately toxic proteins like protease and amylase inhibitors, and lectins are common in plants. Since they are usually associated with seeds and vegetative storage tissues, they are believed to play a role in the protection of those plant parts which are vital for the survival of the species and/or the individual. Moderately toxic lectins (or other proteins) have no immediate effects and therefore cannot protect an individual seed or plant tissue. However, since the delayed reactions brought about in the herbivores eventually result in a reaction of avoidance, the presence of moderately toxic lectins is beneficial for the survival of the plant species. Protection of an