Glycoconjugate Journal 20, 449–460, 2004
© 2004 Kluwer Academic Publishers. Manufactured in The Netherlands.

MINI REVIEW

The identification of inducible cytoplasmic/nuclear carbohydrate-binding proteins urges to develop novel concepts about the role of plant lectins

Els J.M. Van Damme, Nausicaä Lannoo, Elke Fouquaert and Willy J. Peumans

Ghent University, Department of Molecular Biotechnology, Coupure links 653, 9000 Gent, Belgium

During the last few years compelling evidence has been presented for the occurrence of cytoplasmic/nuclear plant lectins that are not detectable in normal plants but are only induced upon application of well-defined stress conditions. Since both the regulation of the expression and the subcellular location indicate that these 'non-classical lectins' are good candidates to play a physiologically important role as mediators of specific protein-carbohydrate-interactions within the plant cell, a critical assessment is made of the impact of these findings on the development of novel concepts about the role of plant lectins. Based on an analysis of the biochemical, molecular and evolutionary data of a jasmonate-induced chitin-binding lectin from tobacco leaves and a salt/jasmonate-induced leaf lectin from rice it is concluded that these lectins most probably interact with endogenous glycans located within the cytoplasmic/nuclear compartment of the plant cell. Several working mechanisms are proposed to explain how these inducible lectins may fulfill an important regulatory or structural role in stressed cells. In addition, the question of the evolutionary relationship(s) between the newly discovered inducible lectins and their 'classical' constitutively expressed homologs is addressed. Evidence is presented that the 'non-classical lectins' represent the main evolutionary line and that some of their corresponding genes were used as templates for genes encoding storage protein-like 'classical' homologs.
Published in 2004.

Keywords: **plant lectin, endogenous role, signal transduction, protein carbohydrate interaction**

Introduction

Many plants contain proteins that bind specifically and reversibly to well-defined sugars. Since a long time, these proteins are usually referred to as 'lectins' or 'agglutinins' because of their ability to clump erythrocytes and/or other types of animal and human cells. Though the terms 'lectin' and 'agglutinin' are still widely used it should be emphasized that agglutination is no longer used as a criterion to consider a protein as a lectin. Along with the advances made in molecular biology in general and plant molecular biology in particular it became evident, indeed, that it is more appropriate to reason in terms of carbohydrate-binding domains, and that accordingly the presence of one or more sugar-binding domains should be the only criterion for a given protein to be classified as a lectin. Based on these considerations plant lectins are defined as plant proteins that possess at least one non-catalytic domain that binds reversibly to a specific mono- or oligosaccharide [1]. A major implication of this definition is that -as earlier proposed for animal lectins [2]-plant lectins do not necessarily consist exclusively of sugar-binding domains but may also comprise an unrelated domain with a totally different structure and biological activity. Evidently, this broadening of the definition has tremendous consequences for the functional study of plant lectins because one has to take into account that the physiological role of some of these proteins depends on two or more domains with a totally different biological activity and that in some instances the key function of the lectin may be determined by the non-lectin domain.

Due to the early discovery in 1888 when Stillmark discovered a proteinaceous agglutinating factor in castor bean seeds plant lectins have a far longer scientific history than most other plant proteins. Moreover, by virtue of their highly specific sugar-binding activity and various biological activities based thereon

To whom correspondence should be addressed: Van Damme Els, Ghent University, Department of Molecular Biotechnology, Coupure links 653, 9000 Gent, Belgium. Tel.: +32 9 264 6086; Fax: +32 9 264 6219; E-mail: ElsJM.VanDamme@UGent.be

(such as cell clumping, mitogenicity, cytotoxicity) plant lectins were not only studied by plant biologists but were the favorite research subjects of scientists working in virtually all possible domains of biological and biomedical sciences. This taken together with the relative ease of purification of reasonable quantities of pure protein preparations eventually explains why plant lectins have so intensively been studied, and as a result are far better characterized with respect to their biochemical and biophysical properties, three dimensional structure and biological activities than any other group of plant proteins. One should realize, indeed, that over 500 different plant lectins have been characterized in some detail for what concerns their biochemical properties and carbohydrate-binding specificity, and that for approximately 200 of these lectins the corresponding genes have been cloned. In addition, a total number of approximately 100 three dimensional structures of plant lectins and lectin/sugar complexes have been resolved [3]. The literature dealing with the sugar-binding activity and biological activities of plant lectins is also very impressive. Though it is evident that not all plant lectins received equal attention there is no doubt that some lectins like concanavalin A, phytohemagglutinin and ricin have been studied in much greater detail than any other plant protein for what concerns their effects on cells and whole organisms. Moreover, lectins are used far more intensively as tools and bioactive proteins in biological and biomedical research than any other group of plant proteins [4–7].

After it became evident that lectins possess carbohydrate-binding activity the search for the physiological role dominated the whole field of plant research for decades. Numerous studies were undertaken to corroborate when and where plants accumulate lectins and how this accumulation can be reconciled with a proposed physiological role. In parallel with this typical plant-physiological approach many efforts were undertaken to find out whether plant lectins possibly interact with other organisms and if so whether this interaction is specific. For a long time, the favorite interacting organisms were the nitrogen fixing bacteria of the genus *Rhizobium* that establish a symbiotic relationship within the root nodules of legumes and to a lesser extent fungi that were believed to be the target organisms of various chitin-binding lectins [1,8–10]. Later, the attention progressively shifted towards phytophagous insects and herbivorous higher animals [1,11]. As a net result, there exists now a huge amount of information about the biology of plant lectins and the effect of these proteins on organisms that interact with plants. Unfortunately, the role of plant lectins is still not fully understood and so the debate on the function of plant lectins either inside or outside the plant is still ongoing. However, one should not be too pessimistic because the apparent lack of insight in the role of plant lectins does not necessarily imply that we know too little about these proteins to answer the question why plants express carbohydrate-binding proteins with a well defined sugar specificity and biological activity. As outlined below, an open-minded and unbiased re-interpretation of the existing data combined with the relevant information deposited in the plant genome/transcriptome databases reveals that the role of most of the 'classical plant lectins' can be fairly well explained in terms of an involvement in protein storage and/or plant defense [1]. However, there is also compelling evidence for the occurrence of 'novel plant lectins' that are involved in specific protein-carbohydrate interactions within the plant cell, and accordingly are essential for the normal growth, development and functioning of the plant.

Some general remarks with respect to the occurrence of plant lectins and their structural and evolutionary relationships

Plants are since a long time considered rich sources of lectins. Though it is true that many common feed and food plants contain reasonable amounts of lectins it is good to realize that the presence of readily detectable quantities of lectin is the exception rather than the rule. At present, lectins have been isolated from approximately 500 different species. Taking into account that the total number of plant species is estimated at several hundred of thousands only a very small part has proven to be a suitable source for the preparation of one or more lectins. Interestingly, these lectins seem to be clustered in some taxonomic groups as is illustrated by the widespread occurrence of abundant agglutinins in *e.g.* legumes, Solanaceae and some families of monocotyledonous plants [3]. However, the occurrence of a lectin in a given plant does not necessarily imply that a similar lectin is present in all closely related species. For example the abundant lectin found in tubers of the Jerusalem artichoke (*Helianthus tuberosus*) could not be detected yet in other *Helianthus* species. Moreover, a recent study of the *Glechoma hederacea* lectin demonstrated that even within a single species the incidence of lectin-negative plants can be reasonably high [12].

The absence of agglutinating activity does not imply that a plant does not express one or more lectins. Analyses of EST databases indicates that orthologs of most of the currently known lectins are expressed in a wide range of plants in which hitherto no lectin activity has been detected. For example, orthologs of jacalin can be retrieved in EST databases of almost all plant species for which a reasonable number of entries have been deposited (W. Peumans, unpublished results). Similarly, a survey of the complete genome indicates that genes encoding orthologs of *e.g.* jacalins and legume lectins are present in *Arabidopsis thaliana*. Furthermore, a search of these databases also reveals the existence of a lot of 'putative lectins' or chimeric proteins with a 'putative lectin domain.' Unfortunately, the carbohydrate-binding activity of these proteins remains to be shown (*e.g.* receptor kinases with a sugar-binding domain) [13–15].

Plant lectins do not represent a single superfamily of plant proteins but definitely exhibit a marked structural diversity. Detailed biochemical, structural and molecular analyses yielded overwhelming evidence that virtually all known plant lectins

Table 1. Overview of the occurrence and cellular localization of plant lectins

Lectin family	*Occurrence*	*Subcellular location*
Amaranthins	Documented only in the genus *Amaranthus* of the Amaranthaceae	Presumably cytoplasmic proteins
Chitin binding lectins comprising hevein domains	Probably ubiquitous in plants	Some proteins are located in the vacuole, others are secreted into the extracellular space
Cucurbitaceae phloem lectins		
Classical Cucurbitaceae phloem lectins	Cucurbitaceae family only	Cytoplasm
Nictaba orthologs	Documented in Solanaceae	Cytoplasm and nucleus
Jacalin-related lectins		
Galactose-specific subgroup	Moraceae family only	Vacuole
Mannose-specific subgroup	Widespread in all major taxonomic groups	Cytoplasm (and nucleus)
Legume lectins	Documented in Fabaceae and Lamiaceae	Vacuole
Monocot mannose-binding lectins	Documented in 8 families of monocots; 3 families of dicots and one liverwort	Vacuole
Type 2 ribosome-inactivating proteins	Documented in >10 plant families from all major taxonomic groups	Vacuole

can be classified into only seven families of structurally and evolutionary related proteins [4], namely the amaranthins, the chitin-binding lectins comprising hevein domain(s), the Cucurbitaceae phloem lectins, the jacalin-related lectins, the legume lectins, the monocot-mannose-binding lectins and the type-2 ribosome-inactivating proteins (Table 1). This classification of lectins in different families is largely based on sequence similarities and structural relationships among lectins. Therefore, lectins for which no sequence information is available cannot be classified. Evidently, one cannot exclude that the few still unclassified lectins belong to different protein families and it may well be possible that novel families of lectins will be discovered in the future. However, it seems very likely that plants - like animals- developed only a limited number of carbohydrate-binding motifs [16–18].

Though seed lectins played an important role in early lectinology plant lectins are certainly not typical seed proteins. A brief survey of the tissue distribution indicates that only a minority of all currently documented plant lectins have been isolated from seeds. Most plant lectins are found, indeed, in vegetative tissues like leaves, stems, bark, fruits, bulbs, tubers, rhizomes and roots [4,8].

In some instances the misconceptions about the tissue distribution of a particular group of plant lectins are so far reaching that they even affect the annotation of these proteins in some databases. This is nicely illustrated by the family of monocot mannose-binding lectins. Both in the NCBI (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and the SMART (http://smart.embl-heidelberg.de/) database this protein family is referred to as 'Bulb-type mannose-specific lectin' (accession numbers gnl|CDD|17765 cd00028 and SM00108, respectively). It is true that the first three dimensional structure for this lectin family was determined by X-ray crystallography analysis of the lectin isolated from bulbs of *Galanthus nivalis* (snowdrop) [19]. However, in the meantime similar lectins have been reported in plants that do not form bulbs such as the orchid *Listera ovata* [20] and the liverwort *Marchantia polymorpha* [21] as well as in the fish *Fugu rupripes* [22]. For these reasons the annotation used in the Pfam database (http://www.sanger.ac.uk/cgi-bin/Pfam/) (Accession number: PF01453; D-mannose binding lectin) is certainly more appropriate.

For a long time the documented occurrence of plant lectins was confined to a handful of legumes and a few other plant species. However, an updated survey reveals that lectins are found in representatives of numerous plant families (Table 1) [3]. Moreover, lectins have not only been isolated from flowering plants but also from at least two lower plants, namely the liverwort *Marchantia polymorpha* [21] and the fern *Phlebodium aureum* [23].

Classical and inducible plant lectins

Until a few years ago virtually all typical plant-related research on lectins was -merely for practical reasons- focused on those carbohydrate-binding proteins that occur in reasonable to high concentrations in seeds or some specialized vegetative tissues. These lectins will further be referred to as 'classical plant lectins.' Thereby the term 'classical' should be interpreted as 'belonging to the group of plant lectins for which there is no obvious direct or indirect evidence that they may be involved in specific lectin-carbohydrate-interactions within the plant cell itself'. The main purpose of the introduction of this term is to distinguish the classical plant lectins from the -hitherto still few- plant lectins that fulfill all the requirements to be considered good candidate mediators of essential

protein-carbohydrate-interactions within the cytoplasmic/nuclear compartment of the plant cell (see below). All lectins with such a potential specific endogenous role are referred to as 'non-classical plant lectins.'

Though the classical plant lectins clearly belong to different families of proteins and, in addition, markedly differ from each other for what concerns the spatial and temporal regulation of their expression, some generalizations can be made with respect to their overall biology. First, all these lectins are fairly to highly abundant proteins representing 0.1–10% of the total protein of the tissue in which they occur. Second, the accumulation of the classical plant lectins is part of a developmental program and is not or only to a very limited extent influenced by external factors other than these affecting normal plant growth and development (*e.g.* day length, season, age). Accordingly, these lectins can certainly not be considered 'inducible' proteins. Third, apart from a few exceptions like amaranthin and the Cucurbitaceae phloem lectins, all the classical plant lectins are synthesized in the endoplasmic reticulum and presumably follow the secretory pathway to reach their final destination [4,8]. According to the available information the majority of these lectins accumulate in the vacuolar compartment but there are also a few documented cases where the lectin is extracellularly secreted into the extracellular space (including the cell wall) [8]. Fourth, specificity studies indicated that most classical plant lectins have a clear preference for 'foreign' over plant specific glycans. This is clearly illustrated by the occurrence of many lectins that bind exclusively to complex animal N- and O-glycans [16].

Based on these considerations the concept was developed that most of the classical plant lectins do not fulfill a specific role in the plant cell itself but rather represent a special class of aspecific defense proteins that help the plant to cope with attacks by herbivorous animals or phytophagous invertebrates, and in case they are not recruited for defense purposes just act as genuine storage proteins [1,4]. In this respect, the classical plant lectins basically differ from animal lectins, because most of these carbohydrate-binding proteins are involved in specific recognition processes within the organism itself [17,24]. Moreover, when involved in defense mechanisms animal lectins do not act directly on the invading agent but indirectly through an involvement in the innate immune system. Evidently, this apparent functional difference raises a fundamental question. If one keeps in mind the increasing evidence that protein-carbohydrate interactions are very important for the normal development and functioning of animal organisms, and the fact that different types of lectins are the mediators of these protein-carbohydrate interactions, it is difficult, indeed, to believe that such interactions are not essential for plants. Therefore, a few years ago we started asking ourselves whether plant lectinologists were possibly looking at the wrong proteins. The basic idea was that it is difficult to envisage how abundant vacuolar or extracellular carbohydrate-binding proteins that are primarily directed against foreign glycans and are constitutively expressed can participate in physiologically important protein-sugar-interactions. If plant lectins with a specific endogenous role exist it is more likely that they are present in physiologically relevant concentrations, and that their expression is induced or enhanced by some specific exogenous or endogenous stimuli (such as plant hormones or biotic/abiotic stress) in those plant tissues or cells that react upon treatment. In animals physiologically relevant protein-carbohydrate-interactions take place outside the cell, at the cell surface (*i.e.* at the outer side of the plasma membrane) and in the cytoplasmic/nuclear compartment. In plants, cell-cell interactions mediated by molecules exposed on the plasma membrane are probably very rare because a thick and rigid cell wall covers the cells. For the same reason, lectin/carbohydrate-interactions in the cell wall and/or extracellular space -if occurring at all- are unlikely to affect basic cellular processes. Accordingly, the cytoplasm and nucleus may be the principal sites of physiologically relevant protein-carbohydrate-interactions in plant cells. Therefore, one can reasonably expect that plant lectins with a specific function are located in the cytoplasmic and/or nuclear compartment and preferentially interact with endogenous plant glycans. Using these criteria of physiologically relevant abundance, inducible expression and cytoplasmic/nuclear localization, two types of lectins were identified that can be considered candidate proteins with a specific physiological role. To distinguish them from the classical plant lectins they are further referred to as 'non-classical plant lectins.'

Examples of non-classical plant lectins

The first evidence for the occurrence of a lectin that is a suitable candidate to mediate essential protein-carbohydrate-interactions in plants followed the isolation and identification of a mannose-binding lectin (called Orysata) that is induced in rice plants by salt-stress, desiccation and the phytohormones jasmonic acid and abscisic acid. Though this protein was already described in 1990 as SalT (a salt-inducible protein) [25] it was only ten years later identified as a mannose-binding lectin [26]. Characterization of the purified protein revealed that Orysata consists of two identical non-covalently associated subunits of 15 kDa that share a high sequence similarity with and exhibit the same overall fold as the previously isolated 'classical plant lectins' belonging to the family of 'jacalin-related lectins.' In spite of these obvious structural similarities the novel rice lectin fundamentally differed from the classical plant lectins with respect to the regulation of its expression. First, Orysata is not constitutively expressed but is synthesized only under specific stress conditions [25,27]. Secondly, even after induction, the expression level of the lectin remains very low ($<$1 mg/kg leaf and stem material) and, in addition, is confined to the roots and sheats of the plant [25]. Hitherto, the intracellular location of Orysata has not been studied but both the absence of a signal peptide in the primary translation product and the documented cytoplasmic/nuclear location of closely related

orthologs in parenchyma cells of *Calystegia sepium* rhizomes [28] and *Morus nigra* bark [29] strongly suggest that the rice lectin is also located in the cytoplasm and/or nucleus. Taken into account the inducibility of the lectin expression by external factors, the very low expression levels and the very specific localization of the lectin both at tissue level and in the cell, one can conclude that Orysata is a good candidate to be considered a protein with a specific physiological function. At present, the exact role of Orysata is not fully understood but it seems likely that this lectin plays a role in plant's response to well-defined stress-factors and according to a recent report also to infection with the pathogenic fungus *Magnaporthe grisea* [30]. Most probably, Orysata is not a unique lectin but rather represents the prototype of a stress/defense-related type of lectin that is common in cereals. Analysis of the EST databases revealed, indeed, that highly similar orthologs are also expressed in maize, wheat and barley. Moreover, experiments with barley demonstrated that seedlings of this cereal accumulate an ortholog upon exposure to intense light [31]. Besides in cereals, ESTs encoding orthologs of Orysata can be found in almost all species for which a reasonable number of sequences have been deposited, including the moss *Physcomitrella patens* (24.8% sequence identity; 59.3% sequence similarity). At present, it is not clear whether all these Orysata orthologs from other species can be considered genuine stress or defense-related proteins. However, the identification of two mannose-specific jacalin-related lectins that are induced in callus of *Helianthus tuberosus* upon treatment with jasmonic acid (and clearly differ from the abundant and constitutively expressed tuber-specific paralogs) [32], and a similar protein (called ipomoelin) that is specifically induced in *Ipomoea batatas* plants in response to wounding and jasmonate [33] indicates that this may well be the case.

A second type of lectin that meets the requirements to be considered a physiologically active carbohydrate-binding plant lectin was recently discovered in leaves of tobacco plants [34]. This lectin (called Nictaba), cannot be detected in untreated plants but is specifically induced by the phytohormone jasmonic acid. Purification of the lectin has shown that it is a dimeric protein composed of two non-covalently associated 19 kDa subunits. Characterization of the protein and cloning of the corresponding gene revealed that Nictaba exhibits an exclusive specificity towards oligomers of *N*-acetylglucosamine and belongs to the same protein family as the Cucurbitaceae phloem lectins [35]. Immunolocalization studies provided unambiguous evidence that Nictaba is exclusively located in the cytoplasm and nucleus of the leaf parenchyma cells, which is in good agreement with the absence of a signal peptide in the lectin cDNA sequence. Though no orthologs of Nictaba have been isolated or identified yet in other plant species, ESTs encoding putative Nictaba-like proteins could be identified in several other species. An important remark to make concerns the obvious structural and evolutionary relationship between Nictaba and the Cucurbitaceae phloem lectins. Sequence comparisons leave no doubt that both types of lectins share a high sequence similarity (32.7% sequence identity; 59.5% sequence similarity). However, there is a major difference because Nictaba and its orthologs lack the characteristic cysteine-rich C-terminal end of the Cucurbitaceae phloem lectins that is believed to be responsible for the covalent interaction (through inter-molecular disulfide bridge formation) of these lectins with the major phloem protein PP1. Apart from this structural dissimilarity, the temporal and spatial regulation of the expression of the Cucurbitaceae phloem lectins is completely different from that of Nictaba. The Cucurbitaceae lectins are constitutively expressed, indeed, whereas Nictaba is only synthesized upon induction with jasmonate. In addition, the Cucurbitaceae phloem lectins are exclusively synthesized in the companion cells (a specialized type of cells in the vascular tissue) [36], whereas Nictaba accumulates in all leaf cells except in the vascular tissue [34].

Mode of action of non-classical plant lectins

The identification of lectins that are at least in principle good candidates to fulfill a specific physiological role is only a first step towards a proof of the importance of specific but essential protein-carbohydrate-interactions in plant cells. Two key questions remain to be answered namely (i) what are the receptor carbohydrates or glycans for these lectins, and (ii) how does an interaction between the lectins and their natural receptors affect or regulate basic processes in the plant cell?

Though Orysata was the first non-classical plant lectin to be identified the question of the possible receptor for this mannose-binding lectin and its mode of action has not been addressed yet. When Orysata was first purified [26] there was still some uncertainty about the subcellular location of this lectin because at that time there was no experimental evidence for the cytoplasmic/nuclear location of mannose-specific jacalin-related lectins. However, as soon as the results of the localization studies of the *Calystegia sepium* [28] and *Morus nigra* lectins [29] provided strong evidence for a cytoplasmic/nuclear location the problem of the possible receptor glycans for Orysata became even more complicated because virtually no information existed about the possible occurrence and *a fortiori* identity of cytoplasmic and/or nuclear mannose-containing glycoconjugates that can act as receptors for this rice lectin.

In this respect, the identification of Nictaba as a cytoplasmic/nuclear lectin with an exclusive specificity towards oligomers of *N*-acetylglucosamine (GlcNAc) was conceptually simpler because the well-documented occurrence of cytoplasmic and nuclear glycoproteins carrying O-linked GlcNAc or GlcNAc-oligomers provided a whole group of potential receptor molecules. It was proposed that the lectin binds to constitutively expressed chito-oligosaccharide-containing glycoconjugates present in the cytoplasm and/or nucleus [34]. It should be mentioned here that in contrast to animal O-GlcNAc glycans, which consist of a single GlcNAc, O-glycans comprising 5 or more GlcNAc residues have been reported in plants [37–38]. As suggested previously, Nictaba can by virtue of its interaction

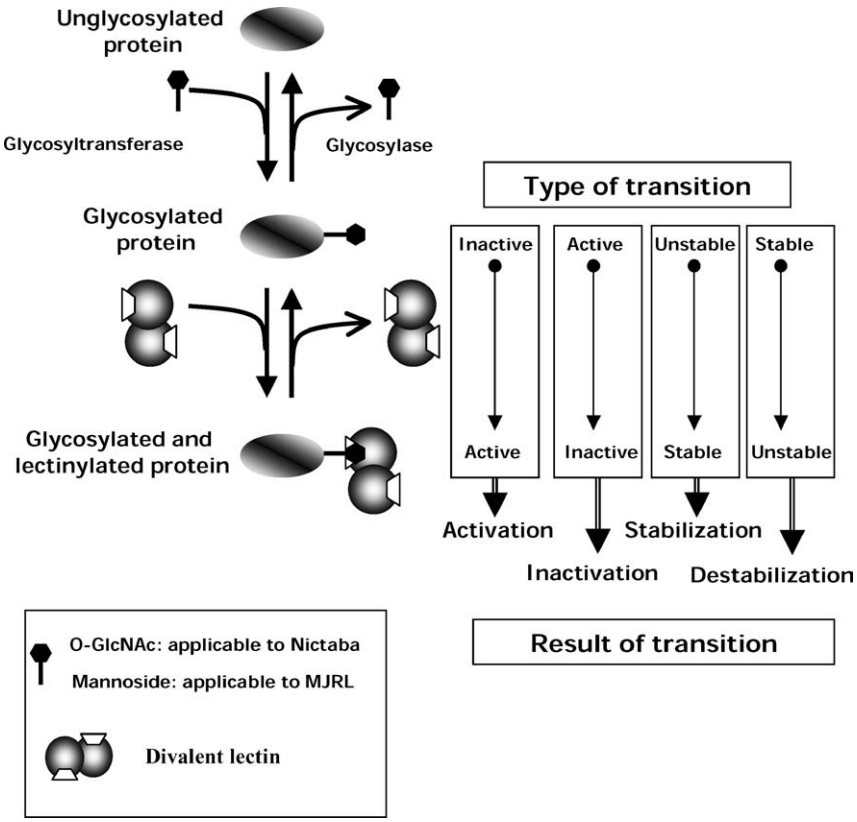

Figure 1. Schematic representation of possible direct effects of the binding of cytoplasmic/nuclear lectins to glycosylated regulatory proteins. Binding of the lectin to a glycosylated protein results in the formation of a lectinylated protein. Depending on the activity/stability-state of the lectinylated and free form, lectinylation results in different effects. Depending on the nature of the glycans (O-GlcNAc and O- or N-mannoside) the scheme is applicable to Nictaba and mannose-specific jacalin-related lectins (MJRL).

with these O-glycans be involved in the regulation of gene expression in stressed plants through a modulation of O-linked N-acetylglucosamine-dependent cell signaling [34]. Based on theoretical considerations the model of the mode of action of Nictaba was further refined. Basically, three different mechanisms can be operative. First, binding of Nictaba to O-linked GlcNAc directly affects the activity and/or stability of soluble regulatory proteins (like transcription and translation factors), protein kinases or receptors in the cytoplasm and/or nucleus (Figure 1). Second, Nictaba can convert physiologically inactive monomeric proteins into activated di- or oligomers (Figure 2). Third, Nictaba can modulate the transport of proteins and RNA between the nucleus and the cytoplasm by narrowing the size of the nuclear pores through binding to the O-(GlcNAc)n glycan chains on the nuclear pore proteins (Figure 3) [34].

In contrast to Nictaba no working hypothesis could be put forward until now to explain the working mechanism of Orysata because no potential receptors have been identified. Though still premature, the occurrence of an apparently widespread family of cytoplasmic/nuclear mannose-binding lectins may indicate that plants possess a signal transduction or regulatory mechanism that relies on mannose-containing receptors. In principle, free cytoplasmic oligomannosides are potential target molecules for the lectins. However, it is difficult to envisage how an interaction with these oligomannosides can generate a specific signal. Therefore, it seems more likely that mannosylated glycoproteins act as receptors for the cytoplasmic/nuclear mannose-binding lectins. Possibly, the plant cell contains soluble mannosylated regulatory proteins in the cytoplasm or nucleoplasm. If so, Orysata and related lectins may act in a similar way as is proposed for Nictaba (Figures 1 and 2). However, a mode of action based on a reversible modulation of the free space of the nuclear pores is rather unlikely because there are no indications that the nuclear pore proteins are mannosylated.

Though there is little doubt that the cytoplasmic/nuclear mannose-binding lectins play an important and specific role in stressed plants, this function must not necessarily rely on an involvement in a signal transduction or regulatory mechanism but can also be based on the formation of lectin-glycoconjugate-complexes with a structural role. Experimental indications for such a structural role have not been obtained yet for a plant lectin, but there is conclusive evidence that in the slime mold *Dictyostelium discoideum* a mannose-binding protein that is closely related to the monocot mannose-binding lectins acts as an important structural element. This lectin, which has been called comitin consists of an N-terminal domain equivalent to the *Galanthus nivalis* agglutinin fused through a short linker

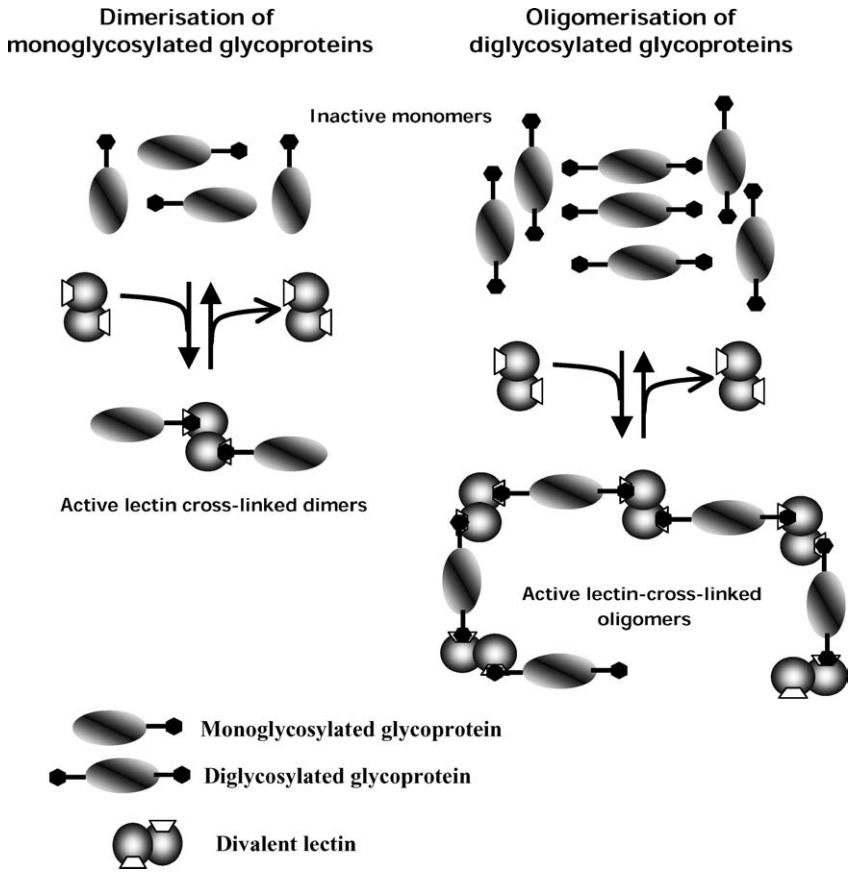

Figure 2. Schematic representation of a reversible cross-linking/oligomerisation of inactive monomeric regulatory proteins by cytoplasmic/nuclear lectins. Cross-linking of two monoglycosylated glycoproteins by a divalent lectin results in dimerization of the glycoproteins. When the protein is substituted by two or more glycans complex oligomers can be formed. Depending on the nature of the glycans (O-GlcNAc and O- or N-mannoside) the scheme is applicable to Nictaba and mannose-specific jacalin-related lectins.

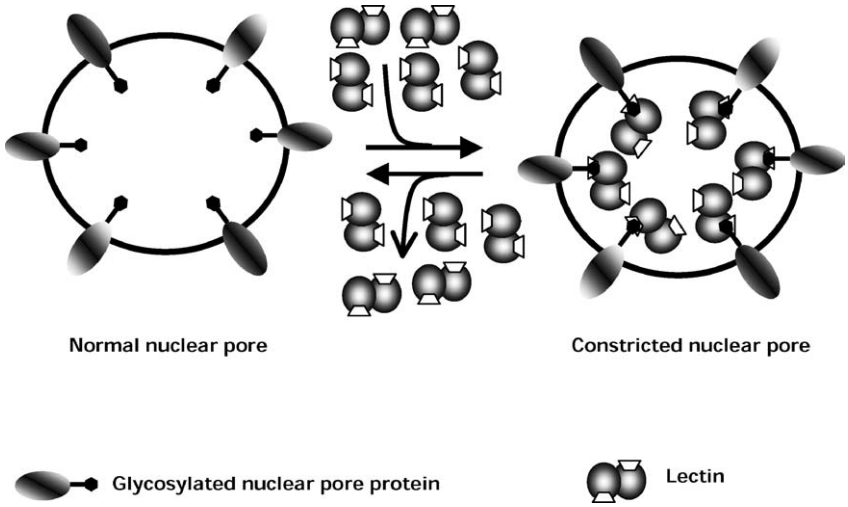

Figure 3. Schematic representation of a reduction of the free space in the nuclear pores by binding of lectins to glycosylated nuclear pore proteins.

to a C-terminal actin-binding domain [39]. It has been demonstrated that comitin binds to vesicle membranes (mainly Golgi vesicles) via its lectin domain and to subdomain 1 of F-actin by its actin-binding domain. As a result, the bifunctional protein is capable of cross-linking Golgi vesicles to the cytoskeleton. The receptor glycans for the lectin domain of comitin have not been identified yet but they correspond most probably to mannose-containing N-glycans located on the cytoplasmic side of glycoproteins residing in the membrane of the Golgi vesicles or to mannosylated glycolipids exposed on the cytoplasmic surface of these vesicles.

What is important from the work with comitin is the demonstration of the occurrence -at least in *Dictyostelium discoideum*- of receptors for a cytoplasmic mannose-binding protein. It is evident that the mode of action of comitin cannot simply be extrapolated to the cytoplasmic/nuclear mannose-specific plant lectins due to the absence of an actin-binding domain. However, if one assumes on the analogy of *Dictyostelium discoideum* that plants also contain mannose-containing glycoconjugates exposed into the cytoplasm, one can start developing models to explain the mode of action of *e.g.* Orysata in establishing key structural elements. A possible simple mechanism could be based on the cross-linking of Golgi vesicles (or possibly other vesicles with surface exposed mannose-residues) by the multivalent lectin (Figure 4). One can imagine, for example, that extensive cross-linking of these vesicles results in the formation of a rigid complex structure that may help to maintain the structural integrity of the cytoplasm and accordingly the whole cell. This could explain why in rice plants Orysata is specifically induced under drought and salt stress (which both have severe consequences for the integrity of the cytoplasm). Similar mechanism may be based on cross-linking of different types of vesicles or on cross-linking between vesicles and other yet unknown structures. Though genome/proteome analyses yielded no evidence for the occurrence of comitin-like proteins in plants, it cannot be precluded that mannose-binding proteins with a homologous function exist in plants (Figure 4). Possibly, the occurrence of chimerolectins comprising a domain similar to Orysata fused to an unrelated domain with a yet unknown activity/function points in that direction. Several

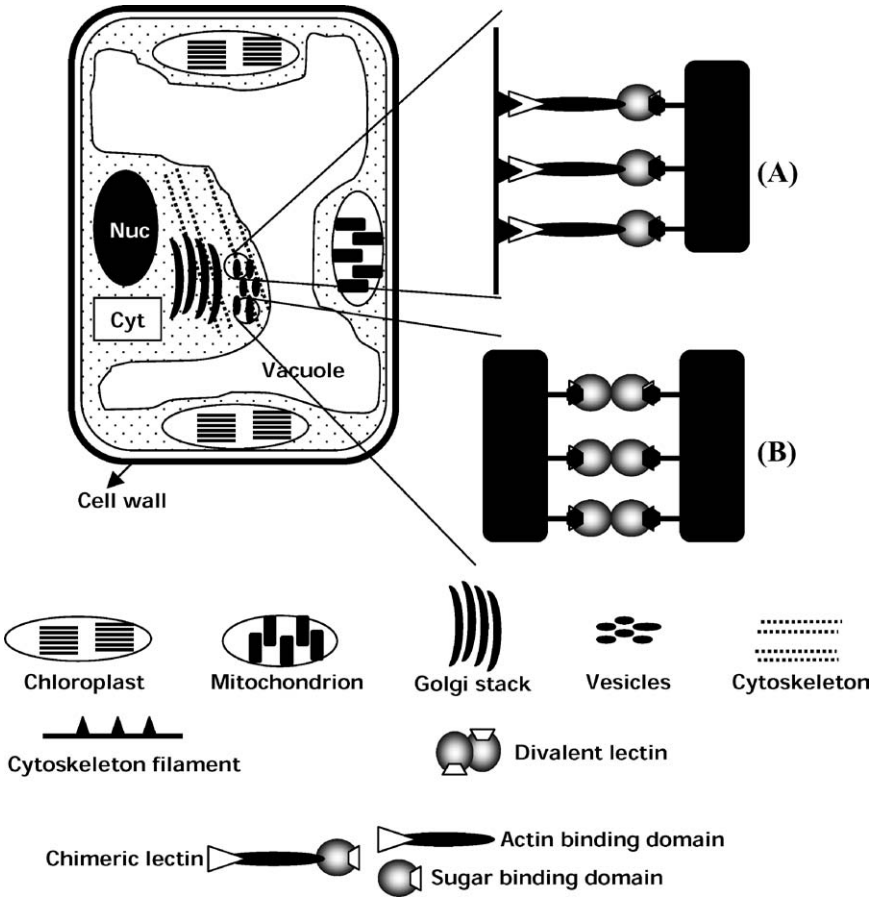

Figure 4. Schematic representation of possible interactions of cytoplasmic mannose-binding lectins in the plant cell. A first type of interaction, mediated by a chimeric lectin consisting of an actin-binding domain and a sugar binding domain, allows a cross-linking between a Golgi-derived vesicle and a cytoskeleton filament (A). In the second type of interaction, a divalent lectin cross-links two Golgi-derived vesicles (B). Nuc and Cyt refer to nucleus and cytoplasm, respectively.

(putative) proteins have been identified indeed, that consist of an N-terminal domain with unknown function linked to a C-terminal domain equivalent to a jacalin domain. Examples are a wheat protein that is specifically induced upon infestation with the hessian white fly [40], a protein that is believed to mediate vernalization-induced flowering in winter wheat [41], and a so-called beta-glucosidase-aggregating factor from maize [42].

Evolutionary and functional relationships between classical and non-classical plant lectins

The occurrence of both 'classical' and 'non-classical' lectins within a single plant lectin family raises the question of the evolutionary and functional relationship between both types of lectins. Since hitherto, non-classical lectins have been identified only within the Cucurbitaceae phloem lectins and the jacalin-related lectins, this issue is for the time being only relevant for these two families.

Due to the detailed biochemical, molecular and structural analyses of a relatively large number of classical jacalin-related lectins and several non-classical orthologs, it was possible to trace the molecular evolution of the whole family and draw important conclusions with respect to the functional specialization of the different groups. As has already been pointed out previously, jacalin-related lectins are fairly widespread over the plant kingdom [28,29]. Representatives of this lectin family have been isolated from numerous flowering plants belonging to different taxonomic groupings and from two lower plants (namely the cycad *Cycas revoluta* [43] and the fern *Phlebodium aureum* [23]). Moreover, analysis of EST databases indicates that orthologs are expressed in almost all species for which a reasonably high number of sequences have been deposited. Sequence comparisons indicate that the non-classical orthologs represent the main evolutionary line, which can be traced from the moss *Physcomitrella patents* throughout all major taxa of modern flowering plants. All these non-classical lectins are cytoplasmic/nuclear proteins that presumably exhibit specificity towards mannose and consist of protomers that are not post-translationally processed. At several occasions, evolutionary events took place whereby a sequence encoding a genuine non-classical lectin was duplicated and acquired a different promoter. The resulting protein was originally identical to the parent non-classical paralog but was no longer inducible but subject to either a constitutive or a developmental regulation. For example, expression of the novel lectins under the control of a promoter that causes a strong expression in developing seeds or vegetative storage organs explains why the very abundant mannose-specific jacalin-related lectins in rhizomes of *Calystegia sepium* and bark of *Morus nigra*, and seeds of the legume tree *Parkia platycephala* behave as genuine vegetative and seed storage proteins, respectively [28,29,44]. Though these storage protein-like jacalin-related lectins resemble the genuine inducible orthologs for what concerns their biosynthesis, molecular structure and subcellular location, there are some marked differences in fine specificity. In general, the exclusive specificity towards mannose/oligomannosides that is very typical for the inducible lectins is less pronounced for the storage protein-like lectins, which usually exhibit a more promiscuous specificity. There is even one documented case in which the specificity towards mannose is almost lost. Detailed specificity studies have demonstrated, indeed, that the *Morus nigra* bark lectin MornigaM, which is structurally similar to the other mannose-specific jacalin-related lectins, accommodates not only mannose but also glucose, galactose and N-acetylglucosamine in its binding site [45]. Structural analyses revealed that this polyspecific character is due to an unusual widening of the binding site. As a result of this altered specificity, the agglutination activity of MornigaM towards animal and human erythrocytes is several orders of magnitude higher than that of any other mannose-specific jacalin-related lectin [46]. Based on this observation, it was proposed that during the evolution of MornigaM from its mannose-specific ancestor, the specificity of the lectin progressively shifted from a preference for typical plant high mannose glycans towards a preference for typical animal glycans, which was interpreted as a strong indication for the involvement of MornigaM in the plant's defense against either phytophagous invertebrates or herbivorous higher animals [29]. Accordingly, MornigaM can be considered the end product of an evolution from a genuine non-classical mannose-specific jacalin-related lectin into a storage protein with a potential role in plant defense. The same holds probably true for the other storage protein-like mannose-specific lectins but in these cases the functional specialization is less obvious from the specificity.

The family of jacalin-related lectins comprises besides the above discussed mannose-specific lectins also a small group of galactose-specific orthologs that comprises jacalin (from jack fruit or *Artocarpus integrifolia*), some virtually identical lectins from other *Artocarpus* species, the *Maclura pomifera* lectin and the galactose-specific *Morus nigra* bark lectin. These galactose-specific lectins differ from the mannose-specific orthologs not only in specificity (galactose versus mannose) but also in subcellular location (vacuolar versus cytoplasmic/nuclear) and biosynthesis (co- and post-translational processing versus no processing). Notwithstanding these differences the mature polypeptides of the mannose-specific and galactose-specific paralogs of *e.g. Morus nigra* still share 60.1% sequence identity and 82.3% similarity at the amino acid level. Based on a detailed analysis of these differences it was proposed that the galactose-specific jacalin-related lectins evolved from a mannose-specific ortholog through the insertion of vacuolar targeting sequences into the primary translation product [28]. Thereby, the 'novel' lectin is not only directed towards a different cell compartment, but also undergoes a post-translational cleavage of the protomer in the direct vicinity of the sugar-binding site. Due to this cleavage, the binding site is far more extended than in the original ancestor lectin and can accommodate different sugars (such as galactose, mannose and glucose). As a result,

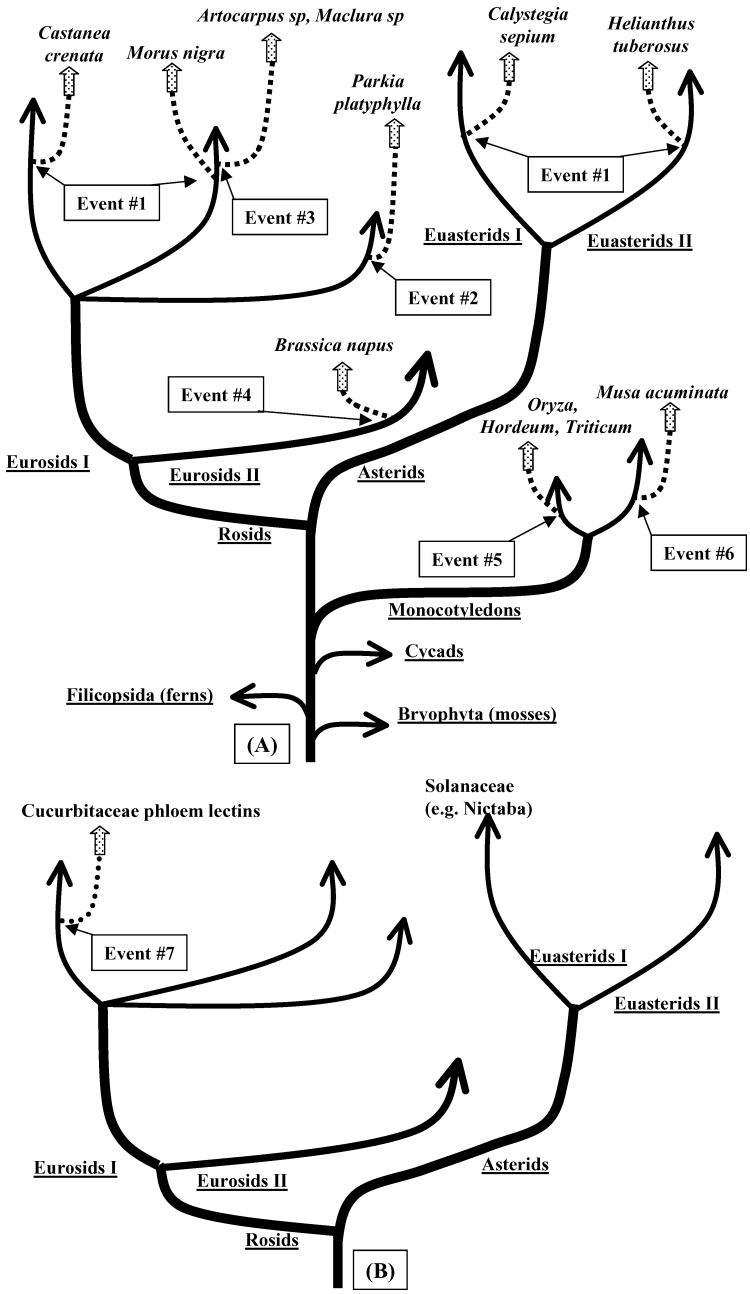

Figure 5. Model of the evolutionary relationship(s) between classical and inducible lectins. Panel (A) gives a general overview of the molecular evolution of the jacalin-related lectins. Full lines represent (the outlines of) the main evolutionary line that comprises the inducible members of this lectin family. The evolutionary events that gave rise to the currently known classical jacalin-related lectins are indicated by dashed lines. Codes for the evolutionary events (boxed) are: Event #1 = gene duplication followed by coupling to a promoter for a vegetative storage protein; Event #2 = gene duplication followed by coupling to a promoter for a seed storage protein; Event #3 = gene duplication followed by the insertion of vacuolar targeting sequences and coupling to a promoter for a seed storage protein; Event #4 = amplification of the lectin domain; Event #5 = gene duplication followed by a fusion to an unrelated domain; Event #6 = gene duplication followed by coupling to a maturation-specific promoter. Underlined terms refer to the main taxonomic groups of modern plants in which jacalin-related lectins have been identified. The taxonomic tree was constructed based on the classification given in the 'Tree of Life Web Project' (http://tolweb.org/tree/). Panel (B) shows an overview of the molecular evolution of Cucurbitaceae phloem lectins. Since only a few members of this lectin family have been identified the scheme is confined to the main taxa of the flowering plants. The presumed main evolutionary line is indicated in full lines. The single evolutionary event that gave rise to the only known group of classical members of this lectin family is indicated by the dashed line. Event #7 stands for gene duplication followed by insertion of a C-terminal peptide (that is believed to form intermolecular disulfide bridges with phloem protein PP1) and coupling to a promoter for phloem specific expression.

jacalin is not a mannose-binding lectin but a polyspecific lectin with a high affinity for the T-antigen [Galβ(1,3)GalNAc] [47]. Taking into account the narrow taxonomic distribution of the galactose-specific jacalin-related lectins (Moraceae family), it has been suggested that the evolutionary event whereby a cytoplasmic mannose-specific lectin was converted into a vacuolar galactose-specific paralog took place at only one occasion in an ancestral Moraceae species. To explain the high expression level (which amounts up to 80% of the total soluble protein in the seeds of jack fruit and 33% in the bark of *Morus nigra*, and is reminiscent to that of genuine seed and bark storage proteins, respectively) it was concluded that the sequence encoding the 'novel' lectin acquired a promoter that ensures a high accumulation rate in the appropriate tissue. Most probably, the change in specificity was not by accident but was intended to confer the novel lectin a specificity that is directed against foreign glycans so that the novel protein could act as an aspecific defense protein against phytophagous invertebrates or herbivorous animals.

In summary, it can be concluded that genes encoding inducible jacalin-related lectins have been used at several occasions as templates for the development of genes encoding storage protein-like classical lectins (Figure 5A). Thereby, the original endogenous function of the lectin was lost, but through more or less dramatic changes in specificity the novel lectin acquired a novel function as an aspecific plant defense protein.

A comparative sequence analysis of the family of the Cucurbitaceae phloem lectins leads to a similar conclusion as for the jacalin-related lectins. According to the available sequence data, putative orthologs of Nictaba are expressed in many species from different taxa [35]. It seems reasonable therefore, that in this family also the non-classical lectins represent the main evolutionary line. Besides Nictaba and closely related orthologs from other species, the genuine Cucurbitaceae phloem lectins are the only other members of this lectin family. Since all currently known phloem-specific lectins from Cucurbitaceae species form a fairly homogeneous group of closely related proteins it seems likely that they all have a common ancestor that probably arose through a single evolutionary event that took place in an early ancestor of the modern Cucurbitaceae. Based on a comparison of the sequences, it is tempting to speculate that the phloem lectins evolved from a Nictaba-like protein through a gene duplication followed by a fusion to a short cysteine-rich domain and the acquisition of a promoter that confines the expression to the companion cells in the vascular tissue. Thereby, the specificity was apparently not affected indicating that the binding site remained well conserved (Figure 5B).

At present, there is no experimental evidence suggesting that the classical lectins of the other lectin families also evolved from non-classical paralogs. However, on the analogy of what can be concluded for the jacalin-related lectins and the Cucurbitaceae phloem lectins there is a reasonable chance that in the (near) future non-classical paralogs will be identified of classical lectins from the other lectin families.

Acknowledgment

This work was supported in part by grants from the Research Council of Ghent University (Grant Nr. 01109203) and Fund for Scientific Research-Flanders (FWO grants G.0113.01 and G.0201.04).

References

1 Peumans WJ, Van Damme EJM, Lectins as plant defense proteins, *Plant Physiol* **109**, 347–52 (1995).
2 Barondes SH, Bifunctional properties of lectins: Lectins redefined, *Trends Biochem Sci* **13**, 480–2 (1988).
3 Van Damme EJM, Peumans WJ, Pusztai A, Bardocz S, *Handbook of Plant Lectins: Properties and Biomedical Applications* (John Wiley & Sons, Chichester, 1998), p. 452.
4 Van Damme EJM, Peumans WJ, Barre A, Rougé P, Plant lectins: A composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles, *Crit Rev Plant Sci* **17**, 575–692 (1998).
5 Lis H, Sharon N, Lectins as molecules and as tools, *Ann Rev Biochem* **55**, 35–67 (1986).
6 Ashraf MT, Khan RH, Mitogenic lectins, *Med Sci Monit* **9**, RA265–9 (2003).
7 Khan F, Khan RH, Sherwani A, Mohmood S, Azfer MA, Lectins as markers for blood group typing, *Med Sci Monit* **8**, RA293–300 (2002).
8 Etzler ME, Plant lectins: Molecular and biological aspects, *Ann Rev Plant Physiol* **36**, 209–34 (1985).
9 Chrispeels MJ, Raikhel NV, Lectins, lectin genes, and their role in plant defense, *Plant Cell* **3**, 1–9 (1991).
10 Hirsch AM, Role of lectins (and rhizobial exopolysaccharides) in legume nodulation, *Curr Opin Plant Biol* **2**, 320–6 (1999).
11 Murdock LL, Shade RE, Lectins and protease inhibitors as plant defenses against insects, *J Agric Food Chem* **50**, 6605–11 (2002).
12 Wang W, Hause B, Peumans WJ, Smagghe G, Mackie A, Fraser R, Van Damme EJM, The Tn-antigen specific lectin from ground ivy is an insecticidal protein with an unusual physiology, *Plant Physiol* **132**, 1322–34 (2003).
13 Barre A, Hervé C, Lescure B, Rougé P, Lectin receptor kinases in plants, *Crit Rev Plant Sci* **21**, 379–99 (2002).
14 Hervé C, Serres J, Dabos P, Canut H, Barre A, Rougé P, Lescure B, Characterization of the Arabidopsis lecRK-a genes: Members of a superfamily encoding putative receptors with an extracellular domain homologous to legume lectins, *Plant Mol Biol* **39**, 671–82 (1999).
15 Hervé C, Dabos P, Galaud JP, Rougé P, Lescure B, Characterization of an Arabidopsis thaliana gene that defines a new class of putative plant receptor kinases with an extracellular lectin-like domain, *J Mol Biol* **258**, 778–88 (1996).
16 Peumans WJ, Barre A, Hao Q, Rougé P, Van Damme EJM, Higher plants developed structurally different motifs to recognize foreign glycans, *Trends Glycosci Glycotechnol* **12**, 83–101 (2000).
17 Kaltner H, Gabius HJ, Animal lectins: From initial description to elaborated structural and functional classification, *Adv. Exp. Med. Biol*. **491**, 79–94 (2001).

18 Kilpatrick DC, *Handbook of Animal Lectins: Properties and Biomedical Applications* (John Wiley & Sons, Chichester, 2000), p. 468.
19 Hester G, Kaku H, Goldstein IJ, Wright CS, Structure of mannose-specific snowdrop (*Galanthus nivalis*) lectin is representative of a new plant lectin family, *Nat Struct Biol* **2**, 472–9 (1995).
20 Van Damme EJM, Smeets K, Torrekens S, Van Leuven F, Peumans WJ, Characterization and molecular cloning of the mannose binding lectins from the Orchidaceae species: *Listera ovata*, *Epipactis helleborine* and *Cymbidium* hybrid, *Eur J Biochem* **221**, 769–77 (1994).
21 Peumans WJ, Barre A, Bras J, Rougé P, Proost P, Van Damme EJM, The liverwort contains a lectin that is structurally and evolutionary related to the monocot mannose-binding lectins, *Plant Physiol* **129**, 1054–65 (2002).
22 Tsutsui S, Tasumi S, Suetake H, Suzuki Y, Lectins homologous to those of monocotyledonous plants in the skin mucus and intestine of pufferfish, *Fugu rubripes*, *J Biol Chem* **278**, 20882–9 (2003).
23 Tateno H, Winter HC, Petryniak J, Goldstein IJ, Purification, characterization, molecular cloning, and expression of novel members of jacalin-related lectins from rhizomes of the true fern *Phlebodium aureum* (L) J. Smith (Polypodiaceae), *J Biol Chem* **278**, 10891–9 (2003).
24 Kuwabara I, Sano H, Liu FT, Functions of galectins in cell adhesion and chemotaxis, *Methods Enzymol* **363**, 532–52 (2003).
25 Claes B, Dekeyser R, Villarroel R, Van den Bulcke M, Bauw G, Van Montagu M, Caplan A, Characterization of a rice gene showing organ-specific expression in response to salt stress and drought, *Plant Cell* **2**, 19–27 (1990).
26 Zhang W, Peumans WJ, Barre A, Houlès-Astoul C, Rovira P, Rougé P, Proost P, Truffa-Bachi P, Jalali AAH, Van Damme EJM, Isolation and characterization of a jacalin-related mannose-binding lectin from salt-stressed rice (*Oryza sativa*) plants, *Planta* **210**, 970–8 (2000).
27 De Souza Filho G, Ferreira BS, Dias JM, Queiroz KS, Branco AT, Bressan-Smith RE, Oliveira JG, Garcia AB, Accumulation of SALT protein in rice plants as a response to environmental stresses, *Plant Sci* **164**, 623–8 (2003).
28 Peumans WJ, Hause B, Van Damme EJM, The galactose-binding and mannose-binding jacalin-related lectins are located in different subcellular compartments, *FEBS Lett* **477**, 186–92 (2000).
29 Van Damme EJM, Hause B, Hu J, Barre A, Rougé P, Proost P, Peumans WJ, Two distinct jacalin-related lectins with a different specificity and subcellular location are major vegetative storage proteins in the bark of the black mulberry tree, *Plant Physiol* **130**, 757–69 (2002).
30 Qin Q-M, Zhang Q, Zhao W-S, Wang Y-Y, Peng Y-L, Identification of a lectin gene induced in rice in response to *Magnaporthe grisea* infection, *Acta Bot Sin* **45**, 76–81 (2003).
31 Potter E, Beator J, Kloppstech K, The expression of mRNAs for light-stress proteins in barley: Inverse relationship of mRNA levels of individual genes within the leaf gradient, *Planta* **199**, 314–20 (1996).
32 Nakagawa R, Yasokawa D, Okumura Y, Nagashima K, Cloning and sequence analysis of cDNA coding for a lectin from *Helianthus tuberosus* callus and its jasmonate-induced expression, *Biosci Biotechnol Biochem* **64**, 1247–54 (2000).
33 Imanishi S, Kito-Nakamura K, Matsuoka K, Morikami A, Nakamura K, A major jasmonate-inducible protein of sweet potato, ipomoelin, is an ABA-independent wound-inducible protein, *Plant Cell Physiol* **38**, 643–52 (1997).
34 Chen Y, Peumans WJ, Hause B, Bras J, Kumar M, Proost P, Barre A, Rougé P, Van Damme EJM, Jasmonic acid methyl ester induces the synthesis of a cytoplasmic/nuclear chitooligosaccharide-binding lectin in tobacco leaves, *FASEB J* **16**, 905–7 (2002).
35 Dinant S, Clark AM, Zhu Y, Vilaine F, Palauqui JC, Kusiak C, Thompson GA, Diversity of the superfamily of phloem lectins (phloem protein 2) in angiosperms, *Plant Physiol* **131**, 114–28 (2003).
36 Bostwick DE, Dannenhoffer JM, Skaggs MI, Lister RM, Larkins BA, Thompson GA, Pumpkin phloem lectin genes are specifically expressed in companion cells, *Plant Cell* **4**, 1539–48 (1992).
37 Heese-Peck A, Cole RN, Borkhsenious ON, Hart GW, Raikhel NV, Plant nuclear pore complex proteins are modified by novel oligosaccharides with terminal N-acetylglucosamine, *Plant Cell* **7**, 1459–71 (1995).
38 Heese-Peck A, Raikhel NV, The nuclear pore complex, *Plant Mol Biol* **38**, 145–62 (1998).
39 Jung E, Fucini P, Stewart M, Noegel AA, Schleicher M, Linking microfilaments to intracellular membranes: The actin-binding and vesicle-associated protein comitin exhibits a mannose-specific lectin activity, *EMBO J* **15**, 1238–46 (1996).
40 Williams CE, Collier CC, Nemacheck JA, Liang C, Cambron SE, A lectin-like wheat gene responds systemically to attempted feeding by avirulent first-instar Hessian fly larvae, *J Chem Ecol* **28**, 1411–28 (2002).
41 Yong WD, Xu YY, Xu WZ, Wang X, Li N, Wu JS, Liang TB, Chong K, Xu ZH, Tan KH, Zhu ZQ, Vernalization-induced flowering in wheat is mediated by a lectin-like gene VER2, *Planta* **217**, 261–70 (2003).
42 Esen A, Blanchard DJ, A specific beta-glucosidase-aggregating factor is responsible for the beta-glucosidase null phenotype in maize, *Plant Physiol* **122**, 563–72 (2000).
43 Yagi F, Iwaya T, Haraguchi T, Goldstein IJ, The lectin from leaves of Japanese cycad, *Cycas revoluta* Thunb. (gymnosperm) is a member of the jacalin-related family, *Eur J Biochem* **269**, 4335–41 (2002).
44 Mann K, Farias CM, Del Sol FG, Santos CF, Grangeiro TB, Nagano CS, Cavada BS, Calvete JJ, The amino-acid sequence of the glucose/mannose-specific lectin isolated from *Parkia platycephala* seeds reveals three tandemly arranged jacalin-related domains, *Eur J Biochem* **268**, 4414–22 (2001).
45 Houlès Astoul C, Peumans WJ, Van Damme EJM, Barre A, Bourne Y, Rougé P, The size, shape and specificity of the sugar-binding site of the jacalin-related lectins is profoundly affected by the proteolytic cleavage of the subunits, *Biochem J* **367**, 817–24 (2002).
46 Rougé P, Peumans WJ, Barre A, Van Damme EJM, A structural basis for the difference in specificity between the two jacalin-related lectins from mulberry (*Morus nigra*) bark, *Biochem Biophys Res Commun* **304**, 91–7 (2003).
47 Bourne Y, Houlès Astoul C, Zamboni V, Peumans WJ, Menu-Bouaouiche L, Van Damme EJM, Barre A, Rougé P, Structural basis for the unusual carbohydrate-binding specificity of jacalin towards galactose and mannose, *Biochem J* **364**, 173–9 (2002).

Received 23 December 2003; revised 10 May 2004;
accepted 11 May 2004