Plant Physiol. (1996) 112: 1531–1540

Structure-Function Relationship of Monocot Mannose-Binding Lectins[1]

Annick Barre, Els J. M. Van Damme, Willy J. Peumans, and Pierre Rougé*

Institut de Pharmacologie et Biologie Structurale, Unité Propre de Recherche Centre National de la Recherche Scientifique No. 9062, Faculté des Sciences Pharmaceutiques, 35 Chemin des Maraîchers, 31062 Toulouse, France (A.B., P.R.); and Laboratory for Phytopathology and Plant Protection, Katholieke Universiteit Leuven, Willem de Croylaan 42, 3001 Leuven, Belgium (E.J.M.V.D., W.J.P.)

The monocot mannose-binding lectins are an extended superfamily of structurally and evolutionarily related proteins, which until now have been isolated from species of the Amaryllidaceae, Alliaceae, Araceae, Orchidaceae, and Liliaceae. To explain the obvious differences in biological activities, the structure-function relationships of the monocot mannose-binding lectins were studied by a combination of glycan-binding studies and molecular modeling using the deduced amino acid sequences of the currently known lectins. Molecular modeling indicated that the number of active mannose-binding sites per monomer varies between three and zero. Since the number of binding sites is fairly well correlated with the binding activity measured by surface plasmon resonance, and is also in good agreement with the results of previous studies of the biological activities of the mannose-binding lectins, molecular modeling is of great value for predicting which lectins are best suited for a particular application.

Plant lectins are an extended group of proteins that, according to a recently updated definition, comprise all plant proteins possessing at least one noncatalytic domain that binds reversibly to specific mono- or oligosaccharides (Peumans and Van Damme, 1995). Due to advances in the biochemistry and molecular biology of plant lectins during the last decade, the structural and evolutionary relationships between the different members of this apparently very heterogeneous group of proteins have become increasingly evident. At present, the majority of all currently known plant lectins can be classified into four groups of evolutionarily related proteins: the legume lectins (Sharon and Lis, 1990), the chitin-binding lectins containing hevein domains (Raikhel and Broekaert, 1993), the type 2 ribosome-inactivating proteins (Barbieri et al., 1993), and the so-called monocot Man-binding lectins. Legume lectins are confined to species of the Leguminoseae, whereas the chitin-binding lectins and type 2 ribosome-inactivating proteins occur in several taxonomically unrelated plant families. Until now, the monocot Man-binding lectins have been found in only five families: Amaryllidaceae, Alliaceae, Araceae, Orchidaceae, and Liliaceae. Extensive studies of numerous monocot Man-binding lectins and molecular cloning of their corresponding genes have shown that they all belong to a single superfamily of evolutionarily related proteins, which not only have a marked sequence homology but also exhibit an exclusive specificity toward Man (Van Damme et al., 1987, 1988, 1991b, 1992a, 1992b, 1993a, 1993b, 1994a, 1994b, 1994c, 1995, 1996a, 1996b; Koike et al., 1995).

The monocot Man-binding lectins have received much recent interest. The exclusive specificity of these lectins toward Man (Shibuya et al., 1988; Kaku et al., 1992; Saito et al., 1993) has been exploited for the analysis and isolation of Man-containing glyconjugates. Other applications in biomedical research are based on the potent inhibitory effect of some monocot Man-binding lectins on human and animal retroviruses (including HIV) (Balzarini et al., 1991, 1992), and on their ability to block the adhesion receptors of Man-fimbriated *Escherichia coli* in the small intestine of rats (Pusztai et al., 1993). Monocot Man-binding lectins have also become an important tool in plant protection and plant biotechnology because their genes confer resistance against sucking insects and nematodes (Hilder et al., 1995). In addition, the particular structure and organization of the

[1] Supported by the Centre National de la Recherche Scientifique and the Conseil Régional de Midi-Pyrénées (A.B., P.R.), the Catholic University of Leuven (OT/94/17), and the National Fund for Scientific Research (Belgium, Fonds voor Geneeskundig Wetenschappelijk Onderzoek grant no. 2.0046.93). W.P. is Research Director and E.V.D. is a postdoctoral fellow of this fund.

* Corresponding author; e-mail rouge@mail.cict.fr; fax 33–61–55–16–72.

Abbreviations: AAA, *Allium ascalonicum* agglutinin; ACA, *Allium cepa* agglutinin; AMADOM1, domain 1 of *Arum maculatum* agglutinin; AMADOM2, domain 2 of *Arum maculatum* agglutinin; APA, *Allium porrum* agglutinin; ASAIDOM1, domain 1 of *Allium sativum* agglutinin I; ASAIDOM2, domain 2 of *Allium sativum* agglutinin I; ASAII, *Allium sativum* agglutinin II; AUAII, *Allium ursinum* agglutinin II; AUAI, *Allium ursinum* agglutinin I; AUAG0, lectin polypeptide composing AUAII; AUAG1, lectin polypeptide composing AUAI; AUAG2, lectin polypeptide composing AUAI; CMA, *Clivia miniata* agglutinin; GNA, *Galanthus nivalis* agglutinin; HHA, *Hippeastrum* hybrid. agglutinin; LOA1, dimeric agglutinin 1 of *Listera ovata*; LOA2, dimeric agglutinin 2 of *Listera ovata*; LOMBP, monomeric Man-binding protein of *Listera ovata*; NPA, *Narcissus pseudonarcissus* agglutinin; PHA-E, *Phaseolus vulgaris* erythroagglutinin; PMA, *Polygonatum multiflorum* agglutinin; TLCIDOM1, domain 1 of *Tulipa* cv Apeldoorn lectin CI; TLCIDOM2, domain 2 of *Tulipa* cv Apeldoorn lectin CI; TLMII, *Tulipa* cv Apeldoorn lectin MII; SPR, surface plasmon resonance.

genes encoding the monocot Man-binding lectins are an interesting research subject for the study of plant genes. Unlike all other plant lectins studied thus far, the monocot Man-binding lectins are encoded by large families of closely related genes (Van Damme et al., 1992a, 1992b, 1993a, 1993b, 1994b, 1994c). Although all monocot Man-binding lectins are very similar at the protein level, there are important differences in the processing and posttranslational modifications of the primary translation products of their genes (Van Damme et al., 1992b, 1993a, 1993b; Smeets et al., 1994).

Despite the obvious structural similarities and sequence homologies between the monocot Man-binding lectins, detailed studies of the carbohydrate-binding specificity and biological activities of different members of this lectin family suggest that there are important differences in the structure of their Man-binding sites. For instance, the fact that some monocot Man-binding lectins exhibit a high antiretroviral activity and others are completely inactive can only be explained by a differential binding to the lectin receptors on the lymphocytes. Since a similar reasoning may hold true for other biological activities, an insight into the structure-function relationships of the monocot Man-binding lectins may be helpful in view of the various potential applications of these proteins. To corroborate the structure-function relationships, the three-dimensional structure of the monocot Man-binding lectins has been modeled using the deduced amino acid sequences of the lectin cDNA clones and the coordinates of the x-ray crystallographic analysis of the snowdrop lectin (Hester et al., 1995). The results of the molecular modeling of the lectins, as well as data obtained from binding studies to various glycoproteins by SPR using a biosensor (BIAcore, Pharmacia), allows precise definition of the structure-function relationships for most monocot Man-binding lectins. Moreover, because the data obtained by these two techniques are in good agreement with the results of previously described specificity and activity studies, molecular modeling can be of great value in predicting which lectins are best suited for a particular biological activity.

MATERIALS AND METHODS

Extraction and Isolation of Lectins

All lectins used in the present study were isolated by affinity chromatography on immobilized Man (see refs. mentioned in Table I).

SPR Analysis

Biospecific-interaction analyses of Man-binding lectins with fetuin, asialofetuin, and PHA-E were performed by SPR using a biosensor (BIAcore, Pharmacia).

Fetuin and asialofetuin were puchased from Sigma. PHA-E was obtained from Vector Laboratories (Burlingame, CA). The sensor chip (CM 5) and all of the chemicals required for the activation of the carboxymethy-

Table I. *Monocot Man-binding proteins from Amaryllidaceae, Alliaceae, Orchidaceae, Araceae, and Liliaceae species*

Species	Molecular Structure[a]	Molecular Mass of Subunits	Antiviral Activity[b]	Accession No.[c]	Protein Code
		kD			
Amaryllidaceae					
Galanthus nivalis	α_4	12.5	+	M55555–M55559	GNA
Narcissus cv's	α_3	12.5	+	M88117–M88123	NPA
Hippeastrum hybrid	α_4	12.5	+	M88124–M88133	HHA
Clivia miniata	α_2	12.5	+	L16511–L16514	CMA
Alliaceae					
Allium sativum					
Lectin I	$(\alpha\beta)$	11.5; 12.5	−	M85174–M85177	ASAI (ASAIDOM1/ASAIDOM2)
Lectin II	α_2	12	−	M85171–M85173	ASAII
Allium ursinum					
Lectin I	$\alpha\beta$	11.5; 12.5	±	L14784–L14785	AUAI (AUAG1/AUAG2)
Lectin II	α_2	12	±	L14783	AUAII (AUAG0)
Allium cepa	α_4	12.5	+	L12171	ACA
Allium ascalonicum	α_2	12.5	+	L12172	AAA
Allium porrum	α_4	13	+	L12173	APA
Orchidaceae					
Listera ovata	α_2	12.5	+	L18894/L18896	LOA
	α	14	−	L18895	LOMBP
Araceae					
Arum maculatum	$(\alpha\beta)_2$	12; 12	−	U12197–U12198	AMA (AMADOM1/AMADOM2)
Liliaceae					
Tulipa cv Apeldoorn	α_2	12	+	U23043–U23044	TLMIII
	c_4 and $(\alpha\beta)_4$[d]	28 and 14; 14	−	U23041–U23042	TLCI (TLCIDOM1/TLCIDOM2)
Polygonatum multiflorum	α_4	14	+	U44775	PMA

[a] $\alpha\beta$, Two different subunits encoded by different genes; $(\alpha\beta)$, two different subunits derived from a single precursor. [b] Antiviral activity of the lectins against retroviruses (HIV-1 and HIV-2), +, active, ±, weakly active, −, inactive. [c] Accession number to the Genbank/EMBL data library. [d] This lectin contains uncleaved (c) and cleaved $(\alpha\beta)$ polypeptides.

lated dextran and the immobilization of glycoproteins (100 mM *N*-hydroxysuccinimide, 400 mM *N*-ethyl-*N'*-[3-dimethylaminopropyl] carbodiimide hydrochloride, and 1 M ethanolamine hydrochloride adjusted to pH 8.5 with NaOH) were obtained from Pharmacia. For immobilization, fetuin, asialofetuin, and PHA-E were used at a concentration of 1 mg mL^{-1} in 5 mM sodium acetate (pH 4.0) buffer. Hepes-buffered saline (10 mM Hepes, pH 7.4, 150 mM NaCl, containing 0.05% BIAcore surfactant P20) used for the biosensor measurements was obtained from Pharmacia. Based on the change in SPR response (expressed in resonance units) as a result of the immobilization of the glycoproteins on the carboxymethylated dextran layer covering the sensor chip, an estimated surface concentration of 10 to 15 ng mm^{-2} of dextran was obtained for the immobilized glycoproteins.

Lectins, used at a constant concentration of 100 $\mu g\ mL^{-1}$ in Hepes-buffered saline, were injected for 5 min onto the glycoprotein-bound surface of the sensor chip at a flow rate of 5 $\mu L\ min^{-1}$. The change in the SPR response (in resonance units) was monitored at 25°C for approximately 9 min. The same glycoprotein sensor chip surface was used repeatedly after removing the remaining immobilized lectin by two successive washes with 10 mM HCl and 10 mM NaOH for 2 min each. Asialofetuin contains six oligosaccharide chains, namely three disaccharides (T antigen) *O*-linked to Thr or Ser residues, and three complex glycans *N*-linked to Asn residues (Shinohara et al., 1994; Dill and Olson, 1995) (Fig. 1). In fetuin, the exposed Gal residues of both *O*-linked and *N*-linked saccharides are masked by sialic acid residues. PHA-E contains an *N*-linked high-Man-type glycan with exposed Manα1→2-linked residues and a Xyl-containing oligosaccharide with a Fucα1→3 branched residue (Sturm et al., 1992) (Fig. 1). For inhibition assays, Man used at concentrations ranging from 5 to 100 mM in Hepes-buffered saline was injected at the beginning of the dissociation phase for 5 min at a flow rate of 5 $\mu L\ min^{-1}$, and the change of the SPR response (in resonance units) was monitored at 25°C for approximately 9 min.

Hydrophobic Cluster Analysis

The hydrophobic cluster analysis (Gaboriaud et al., 1987; Lemesle-Varloot et al., 1990) was performed to delineate the structurally conserved β sheets along the amino acid sequences of the investigated lectins and the model lectin from snowdrop (GNA). Hydrophobic cluster analysis plots were generated on a Macintosh LC using the program Hydrophobic Cluster Analysis-Plot2 (Doriane, Paris, France).

Molecular Modeling

Molecular modeling of lectins was carried out on a workstation (Iris 4D25G, Silicon Graphics, Mountain View, CA) using the programs InsightII, Homology, and Discover (Biosym Technologies, San Diego, CA). The atomic coordinates of GNA (code 1MSA) were taken from the Brookhaven Protein Data Bank (Hester et al., 1995) and used to build the three-dimensional models of other lectins. Energy minimization and relaxation of loop regions were carried out by several cycles of steepest descent and conjugate gradient using the Consistent Valence Forcefield of the Discover program. The program Turbo-Frodo (Bio-Graphics, Marseille, France) was run on a workstation (Indigo R3000, Silicon Graphics) to perform the superposition of the models and the docking of Man into the Man-binding sites of the lectins. The program Cameleon (Oxford Molecular, Palaiseau, France) run on the workstation (Indigo R3000) was used to predict the exposure of the putative *N*-glycosylation sites occurring along the amino acid sequences of lectins using various algorithms (Janin, 1979; Hopp and Woods, 1981; Karplus and Schulz, 1985; Kyte and Doolittle, 1982; Parker et al., 1986; Thornton et al., 1986).

Phylogenetic Analysis

The amino acid sequence alignments were performed on a MicroVAX 3100 (Digital, Evry, France) using the ialign program (PIR/NBRF, Washington, DC). The MacClade program (Maddison and Maddison, 1992) was run on a Macintosh LC 630 to build a parsimony phylogenetic tree relating the different monocot Man-binding lectins.

RESULTS AND DISCUSSION

Overview of the Monocot Man-Binding Lectins

To corroborate the structure-function relationships of the monocot Man-binding proteins, we attempted to correlate the modeled three-dimensional structure of the subunits of different members of this superfamily of lectins to their carbohydrate-binding and other biological activities. Because the superfamily of monocot Man-binding lectins is

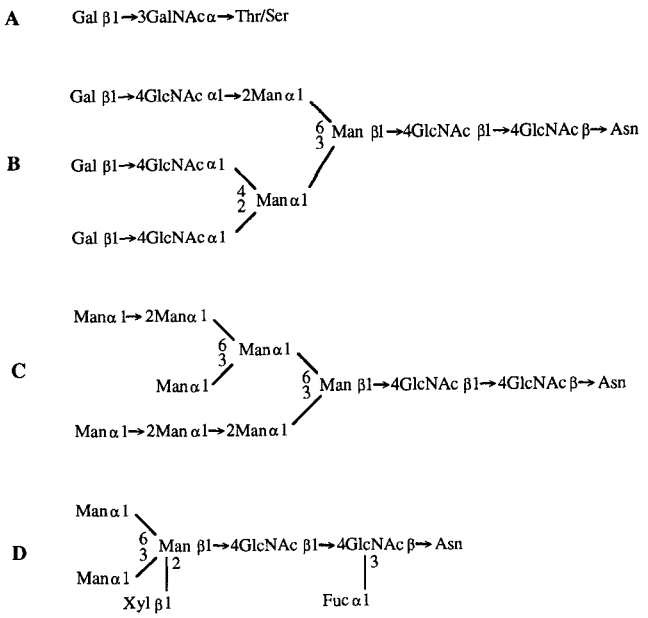

Figure 1. *O*-linked and *N*-linked oligosaccharides found in asialofetuin (A and B) and PHA-E (C and D). In fetuin, exposed Manα1→2 shown in A and B are masked by sialic acid residues.

quite extended, and because there are important differences in the molecular structure (in terms of subunit composition) of the different subgroups, the lectins used in this study and their relevant properties are briefly reviewed in Table I. As shown there, four different molecular forms occur. Most of the lectins are homo-oligomers composed of two, three, or four identical subunits of about 12 kD, which are synthesized as separate polypeptides. It should be mentioned, however, that these homo-oligomeric lectins are synthesized as preproproteins, which are converted into the mature lectin polypeptides by the co-translational cleavage of a signal peptide and the posttranslational removal of a C-terminal peptide (Van Damme and Peumans, 1988; Van Damme et al., 1991b; Hester et al., 1995). Aside from the homo-oligomeric lectins, there are several types of hetero-oligomeric forms. First, there are the heterodimers that result from a noncovalent association between two different (but highly homologous) subunits of about 12 kD, both of which are derived from separate preproproteins that undergo a processing similar to that of precursors of the homo-oligomeric lectins (e.g. *Allium ursinum* lectin I) (Van Damme et al., 1993b). A second type are the heterodimers or heterotetramers, which are composed of two different types of subunits that are derived (through a complex posttranslational processing) from a single precursor with two distinct lectin domains. In some cases the two subunit types are highly homologous (e.g. *Allium sativum* lectin I) (Van Damme et al., 1992b); in others the homology between the two domains is much lower (e.g. *Arum maculatum* lectin) (Van Damme et al., 1995). Finally, the tulip lectin TLCI represents a unique protein. This lectin is a tetramer of four identical subunits of 28 kD containing two separate domains. Since most of the 28-kD polypeptides are cleaved into two smaller subunits, it appears that the native lectin behaves as a hetero-octamer (Van Damme et al., 1996b).

It should be emphasized that the molecular modeling of the different lectins has been done with the sequences of the mature lectin subunits (i.e. after cleavage of the signal peptide, the C-terminal peptide, and, if applicable, the intervening sequence between the two lectin domains of the precursor).

Molecular Modeling of the Monocot Man-Binding Lectins

Biochemical and molecular analyses have shown that all currently known monocot Man-binding lectins are structurally related to each other. Since the structure of the first isolated member of this superfamily of lectins, GNA, has recently been resolved by x-ray crystallography, the coordinates of the latter lectin can now be used for the molecular modeling of the other monocot Man-binding lectins.

As shown by Hester et al. (1995), the three-dimensional structure of GNA corresponds to a β barrel built up of three antiparallel, four-stranded β sheets (subdomains) interconnected by Ω loops (Hester et al., 1995) (Fig. 2). Two Cys residues, Cys^{29} and Cys^{52}, are linked through a disulfide bridge. Each lectin monomer possesses three identical Man-binding sites made of four amino acid residues: Gln, Asp, Asn, and Tyr (Gln^{89}, Asp^{91}, Asn^{93}, and Tyr^{97} for subdomain I; Gln^{57}, Asp^{59}, Asn^{61}, and Tyr^{65} for subdomain II; and Gln^{26}, Asp^{28}, Asn^{30}, and Tyr^{34} for subdomain III) that bind O2 (Asp and Asn), O3 (Gln), and O4 (Tyr) of Man through a network of four hydrogen bonds (see Fig. 5A). Another hydrophobic residue, Val^{95} (subdomain I), Val^{63} (subdomain II), and Val^{32} (subdomain III), interacts with C-3 and C-4 of Man through hydrophobic interactions. Four GNA monomers build a tetramer with 12 well-exposed Man-binding sites.

The high percentages of both identity and similarity observed between the amino acid sequences of GNA (Van Damme et al., 1991b) and other lectins from the families Amaryllidaceae, Alliaceae, Orchidaceae, Araceae, and Liliaceae (Van Damme et al., 1992a, 1992b, 1993a, 1993b, 1994a, 1994b, 1994c, 1995, 1996a, 1996b) (Fig. 3) suggest that the subunits of all of these lectins share a common three-dimensional structure. Furthermore, the hydrophobic cluster analysis plots of the Man-binding lectins look very similar to those of GNA and allow the exact delineation of the 12 strands of antiparallel β sheet found in GNA along the amino acid sequences of these lectins. Although the sequence homologies of the Man-binding lectins from the Liliaceae and Araceae families to the snowdrop lectin are rather low and disturbed by some deletions or insertions occurring along the amino acid sequences, their hydropho-

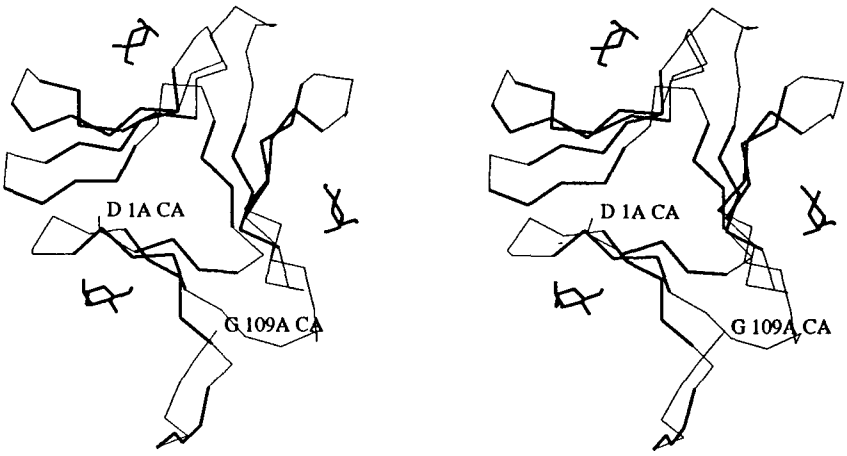

Figure 2. Stereoviews of the α-carbon tracings of GNA. Thick segments correspond to the 12 strands of the β sheet. The Man residues occupying the three Man-binding sites of GNA are represented.

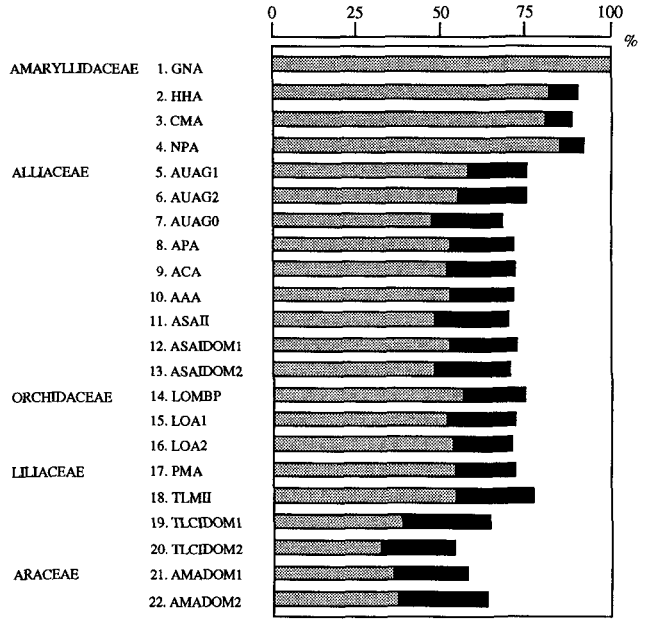

Figure 3. Identity (gray rods) and homology (black rods) percentages relating various monocot Man-binding lectins to GNA taken as reference.

bic cluster analysis plots clearly show that they are structurally related to GNA. Accordingly, the structurally conserved regions common to GNA are easily recognized along their amino acid sequences and can be used to build the three-dimensional models of these lectin monomers using molecular modeling.

In general, the overall three-dimensional conformation of the lectin subunits, built from the x-ray coordinates of GNA, strongly resembles that of GNA except that in some lectins deletions or insertions occur in the loops interconnecting the antiparallel strands of the β sheet (Fig. 4). A closer examination of the modeled structures reveals that the amino acid residues forming the Man-binding sites of GNA have been conserved in the polypeptides of all Amaryllidaceae and Orchidaceae lectins. As a result, the Man-binding sites of these lectins are readily superimposable to those of GNA (Fig. 5). The same holds true for the Alliaceae lectins except for AUAII and ASAIDOM2. In contrast, the amino acid residues that form the Man-binding sites have not been conserved in many of the subunits of the Liliaceae and Araceae lectins. Therefore, the Man-binding sites of these lectins have an altered three-dimensional structure. As a result, the hydrogen bonding scheme of Man is altered in such a way that the binding sites become inactive because of steric clashes due to the replacement of small residues by more extended residues (e.g. the replacement of Asn^{93} by Arg^{93} in the site of subdomain I of AMADOM1 induces a steric hindrance), or because of the disability of some altered residues to participate in hydrogen bonds (e.g. the replacement of Asn^{93} by Leu^{93} in the site of subdomain I of AMADOM2 abolishes one of the two hydrogen bonds connecting O2 of Man to the binding site) (Fig. 5B). Following the same reasoning, one (ASAIDOM2 and PMA), two (TLCIDOM1, AMADOM1, and AMADOM2), or all three (AUAII and TLCIDOM2) Man-binding sites of some lectins cannot accommodate Man. As a result, AUAII and TLCIDOM2 are devoid of Man-binding activity, whereas ASAIDOM2, PMA, TLCIDOM1, AMADOM1, and AMADOM2 have a reduced Man-binding activity.

All monocot Man-binding lectins are encoded by multiple genes. A comparison of the deduced amino acid sequences of different cDNA clones encoding isolectins from snowdrop (*Galanthus nivalis*) (Van Damme et al., 1991a), daffodil (*Narcissus pseudonarcissus*), and amaryllis (*Hippeastrum* spp.) (Van Damme et al., 1992a) revealed that the residues that form the Man-binding site have been conserved and that all of the lectin subunits possess three functional binding sites. However, in the case of garlic (*Allium sativum*) (Van Damme et al., 1992b), some isoforms apparently underwent an amino acid substitution at some positions of residues involved in the sugar-binding sites. As a result, some of the lectin isoforms of ASAIDOM2 and ASAII have only two instead of three active binding sites. It can be expected, therefore, that the latter isoforms exhibit a reduced affinity for Man-containing glycoconjugates.

The noncovalent association of GNA monomers into homotetrameric native lectin molecules containing 12 Man-binding sites is mediated by both hydrogen bonds and hydrophobic interactions interconnecting monomers A, B, C, and D. In fact, two homodimers, A-D and B-C, result from the noncovalent binding of two monomers via hydrogen bonds and van der Waals contacts. These hydrogen bonds occur between the C-terminal end of each monomer (residues Arg^{101}, Trp^{102}, Thr^{104}, and Thr^{108} of monomer A or D and monomer B or C) and three residues (Asp^{91}, Val^{94},

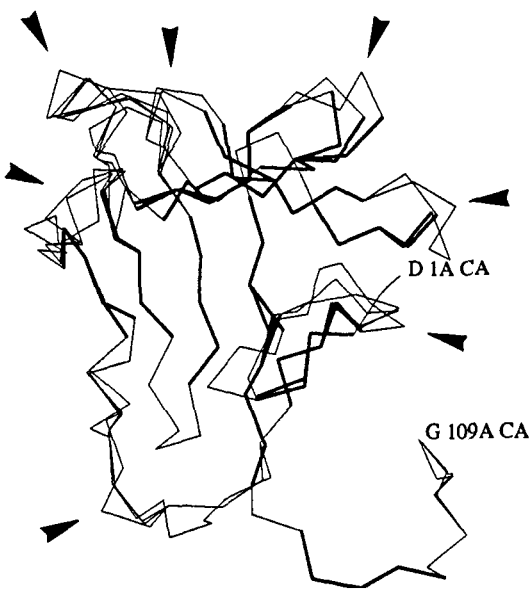

Figure 4. Superposition of the α-carbon tracings of the three-dimensional models of 10 Man-binding lectin monomers (NPA, HHA, AUAG0, AUAG1, ASAII, TLCIDOM1, TLCIDOM2, AMADOM1, AMADOM2, and PMA) built from the coordinates of GNA to the three-dimensional model of the snowdrop lectin. Strands of the β sheet (thick lines) are well superimposed, whereas most of the conformational changes (arrowheads) occur in loops (thin lines).

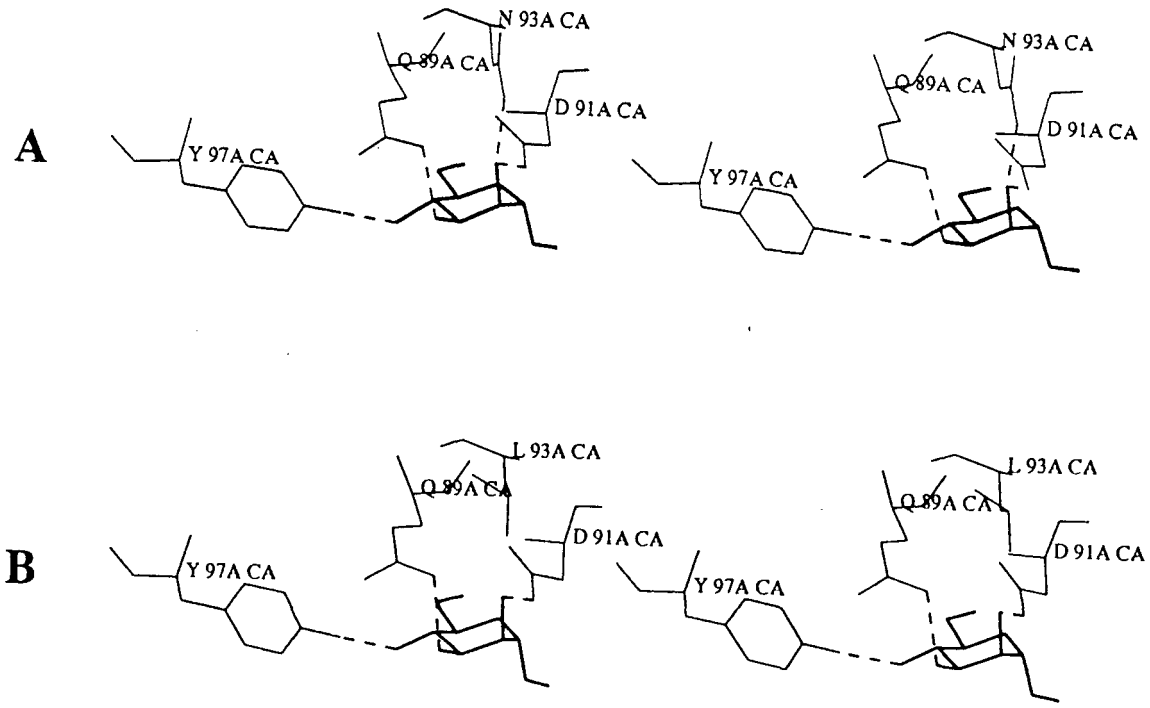

Figure 5. Stereoview showing the docking of Man into the site of subdomain I of GNA (A) and AMADOM2 (B). The replacement of Asn^{93} of GNA by Leu^{93} in the site of subdomain I of AMADOM2 abolishes one of the two hydrogen bonds connecting O2 of Man to the binding site. Therefore, this modified Man-binding site is believed to be either weakly reactive or inactive toward Man.

and Ile^{96}) belonging to the other monomer (D or A and C or B). Two interfaces consisting of van der Waals contacts occurring between two loops (residues 16–20 and 34–38) belonging to monomers A and B and C and D, respectively, interconnect the two homodimers to form a homotetramer A-D/B-C. This homotetramer is arranged as a 12-stranded β barrel and exhibits a wide, central solvent channel rich in hydrophilic residues (Fig. 6). Most of the residues responsible for the association of the GNA monomers in homodimers are conserved in other Man-binding lectins, suggesting that many Man-binding monocot lectins can occur as dimers and tetramers as well. Along this line, the two loops (residues 16–20 and 34–38 in GNA) possibly involved in the connection of dimers in other monocot lectins are less modified by insertions, deletions, or amino acid changes than other loops present on the surface of the lectin monomers. However, the deletion of four residues (36–39) occurring in one of the two loops (residues 34–38) in AMADOM1 (*Arum maculatum*) could prevent dimers to form a tetrameric lectin. Similarly, an insertion of three residues occurring between residues 18 and 19 of the loop corresponding to residues 16 to 20 should prevent the formation of tetramers as predicted by molecular modeling. In fact, the above-mentioned predictions dealing with monomer interactions have to be interpretated with extreme caution, since the native lectin present in *A. maculatum* has an apparent molecular mass of 50 kD, which corresponds to a heterotetramer (Van Damme et al., 1995).

Reactivity of Different Monocot Man-Binding Lectins with Complex Glycan Chains of Fetuin, Asialofetuin, and PHA-E

To correlate the results of the molecular modeling of the monocot Man-binding lectins with their binding specificity and activity, an analysis was made of the interactions of the lectins with the complex glycan chains of fetuin, asialofe-

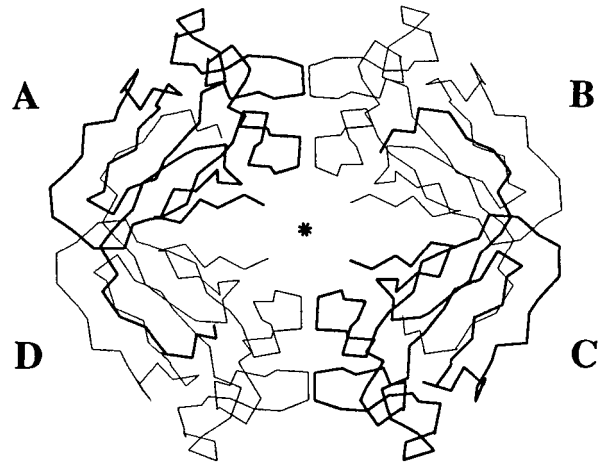

Figure 6. GNA homotetramer showing the noncovalent assembly of monomers A, B, C, and D into a 12-stranded β barrel exhibiting a wide, central solvent channel (*) rich in hydrophilic residues. Each monomer contains three exposed Man-binding sites, resulting in 12 functional sites per tetramer.

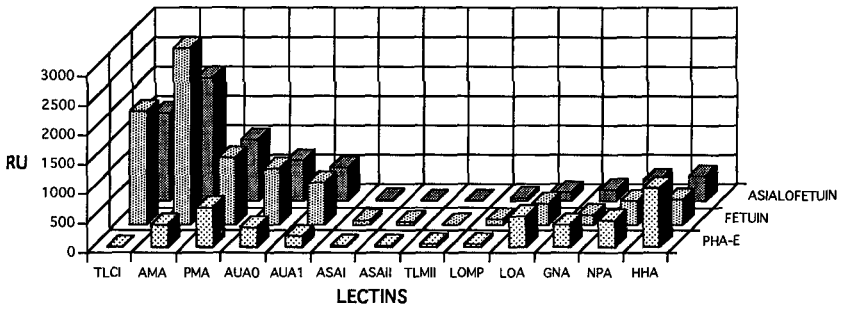

Figure 7. Interaction measured by SPR between 11 lectins and 3 immobilized glycoproteins (asialofetuin, fetuin, and PHA-E). The amount of lectin bound to the immobilized glycoproteins after approximately 9 min (see "Materials and Methods") is expressed in resonance units (RU).

tuin, and PHA-E by SPR. As shown in Figure 7, the monocot Man-binding lectins can be divided into three groups on the basis of their reactivity toward these complex oligosaccharides. The lectins from *A. maculatum*, *Polygonatum multiflorum*, and *Allium ursinum*, as well as the lectin TLCI from *Tulipa*, strongly react with fetuin and asialofetuin, but exhibit a low reactivity toward PHA-E (Fig. 7). Amaryllidaceae and Orchidaceae lectins react moderately with PHA-E and only weakly with fetuin and asialofetuin. Finally, the garlic lectins ASAI and ASAII, the monomeric Man-binding protein from *Listera ovata*, and the lectin TLMII from *Tulipa* are virtually inactive with all three glycoproteins tested.

Although its significance remains unclear, the reactivity of the lectins toward PHA-E is reasonably well correlated with their antiviral activity against human and animal retrovitruses. This presumably depends on the ability of monocot lectins to specifically recognize oligosaccharides shared by both PHA-E and retrovirus surfaces. Previous studies have demonstrated that GNA, NPA, HHA, LOA (Balzarini et al., 1991, 1992), and PMA (Van Damme et al., 1996a) are potent antiviral agents, whereas AUAI and AUAII are only weakly active and ASAI, ASAII, LOMBP, and AMA are completely inactive.

The high reactivity of AMA toward fetuin and asialofetuin is in good agreement with the fact that this lectin can be isolated by affinity chromatography on immobilized fetuin. It is difficult, however, to explain why this lectin, which has only one active Man-binding site per subunit, is so reactive with fetuin and asialofetuin. A possible explanation might be that the second domain of this lectin (which has only a low sequence homology with the other Man-binding lectins) possesses one or more binding sites that lost their ability to bind Man but are capable of binding to another sugar. It is worth mentioning in this context that the second domain of TLCI, which as shown in Figure 8 is the closest relative of the second domain of AMA, exhibits *N*-acetylgalactosamine-binding activity (Van Damme et al., 1996b).

Phylogeny and Molecular Evolution of the Monocot Man-Binding Lectins

The availability of numerous amino acid sequences encoding Man-binding lectins from Amaryllidaceae, Alliaceae, Orchidaceae, Araceae, and Liliaceae species enabled us to analyze the homology and molecular evolution of the different members of this superfamily of monocot Man-binding lectins. According to the classification of the families of the monocots as proposed by Dahlgren et al. (1985), the Alliaceae and Amaryllidaceae are considered to be two closely related families in the order Asparagales. The Liliaceae and Orchidaceae are both classified in the order Liliales, which, like the Asparagales, belongs to the superorder Liliiflorae. In contrast, the Araceae is classified in the order Arales of the superorder Ariflorae. As shown in Figure 8, all Amaryllidaceae, Alliaceae, and Orchidaceae lectins are clustered, although in separate groups, which is in perfect agreement with the above-described taxonomical treatment of the monocots. However, the situation is less clear-cut for the Araceae and Liliaceae lectins. Although these two families are not closely related taxonomically, the sequences of the lectins from *Tulipa* hybrids, *Polygonatum multiflorum*, and *Arum maculatum* are clearly clustered. In addition, it is also striking that all of the monocot Man-binding lectins, which are synthesized as large precursors with two clearly different lectin domains, are found in this cluster.

Irrespective of the discrepancies between our results and the currently accepted classification of the five aforementioned plant families, the phylogenetic tree shown in Figure 8 is in good agreement with the grouping of the monocot Man-binding lectins into a single superfamily of closely related proteins according to both their structural features and their functional properties. By plotting the number of amino acid changes occurring at each residue position

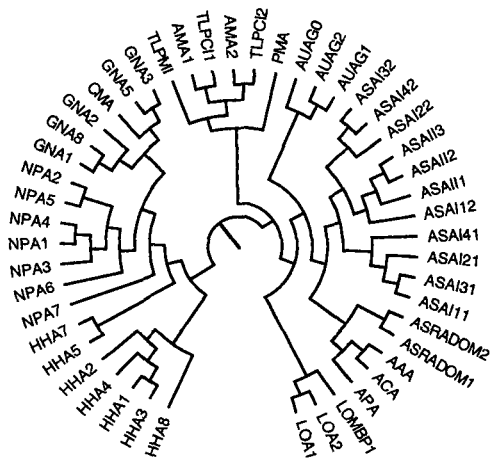

Figure 8. Phylogenetic tree built from the amino acid sequences encoding monocot Man-binding proteins.

along the amino acid sequences of the monocot lectins (Fig. 9), the three Man-binding sites of these lectins clearly appear as differently conserved. The Man-binding site of subdomain III underwent few changes and retains its Man-binding activity in most of the lectins studied. Conversely, the Man-binding sites of subdomains I and II exhibit many more changes, and many are devoid of Man-binding capacity, especially in the Araceae and Liliaceae lectins (see Fig. 9).

CONCLUSION

Monocot Man-binding lectins have been studied in detail at the biochemical and molecular levels in representative species of five different families. Therefore, these lectins represent a unique model system for a study of the structure-function relationships of carbohydrate proteins from a large taxonomic group. Using a combination of molecular modeling and carbohydrate-binding studies, interesting correlations have been found between the amino acid sequence, three-dimensional structure, and sugar-binding activity of the different lectins.

Molecular modeling clearly indicated that, in spite of differences in amino acid sequences, the subunits of all monocot Man-binding lectins share a common three-dimensional structure similar to that of GNA. These proteins exhibit a β-barrel structure made of three subdomains, each consisting of four strands of antiparallel β sheet interconnected by Ω loops. Most of the residue changes, including deletions or insertions, essentially occur in exposed loops. As a result, some of these Man-binding sites homologous to those present in the GNA monomer are predicted to lose their Man-binding capacity. Accordingly, in addition to fully reactive lectin monomers that possess three active Man-binding sites (CMA, NPA, HHA, AUAG1, AUAG2, APA, ACA, AAA, ASAII, ASAIDOM1, TLMII, LOMBP, LOA1, and LOA2), other monomers exhibit only two (ASAIDOM2, PMA), one (TLCIDOM1, AMADOM1, and AMADOM2), or no (AUAG0 and TLCIDOM2) active sugar-binding sites.

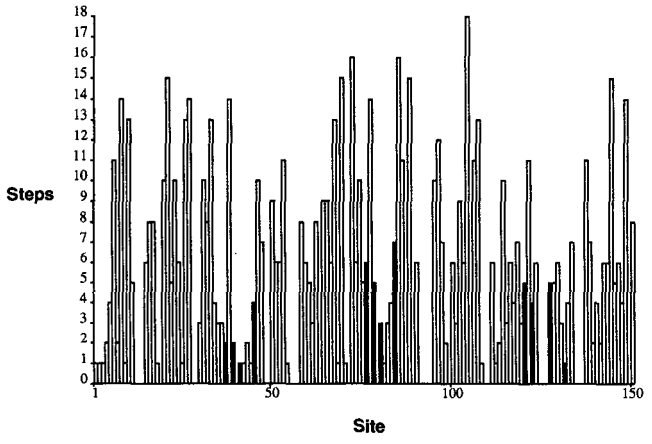

Figure 9. Plot of the amino acid changes occurring at each position (site) along the amino acid sequences of monocot Man-binding lectins. Black rods correspond to residues forming the three Man-binding sites of subdomains III, I, and II, respectively.

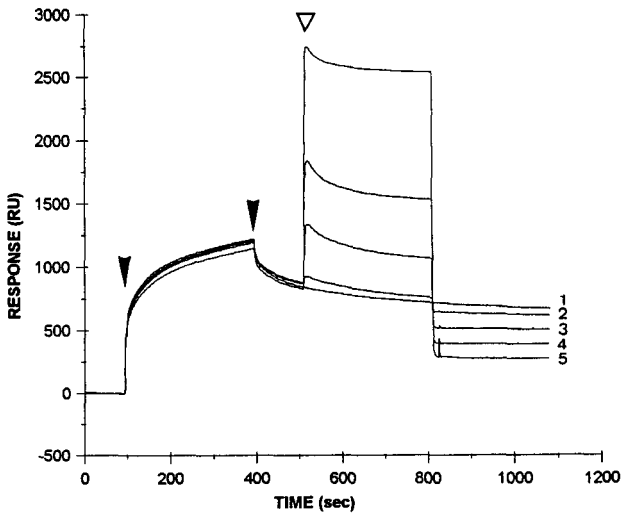

Figure 10. Effect of Man added at the beginning of the dissociation phase (white arrowhead) at concentrations of 5 mM (2), 25 mM (3), 50 mM (4), and 10 mM (5) on the interaction between HHA and surface-bound calf asialofetuin (1 = no Man added). HHA was injected at a concentration of 100 μg mL^{-1}; black arrowheads indicate the beginning and end of the injections. Whatever the concentration used, Man does not totally dissociate the HHA-asialofetuin complex.

Although studies on the interactions of the lectins with various glycoproteins containing exposed Gal (fetuin and asialofetuin) or Man residues (PHA-E) are in good agreement with the predictions derived from the molecular modeling studies, there is one exception. AUAG0 (AUAII), which on the basis of the modeling studies possesses no active Man-binding site, reacts with PHA-E and, more strongly, with both fetuin and asialofetuin. These discrepancies could result from the fact that the Man-binding site of the monocot lectins responsible for their binding to monosaccharides is part of a more extended site, which could account for stronger binding to complex glycans. In this respect, α1,6-linked mannotriose was reported to be 10 to 20 times more active than Man in inhibiting the daffodil (NPA) and amaryllis (HHA) lectins (Kaku et al., 1990). Similarly, inhibition experiments carried out in the presence of Man as the inhibitor at concentrations ranging from 5 to 100 mM showed that Man added during the dissociation phase was unable to totally reverse the HHA-asialofetuin interaction (Fig. 10). These results indicate that the affinity of monocot lectins toward complex glycans is higher than that for simple sugars. Lectins from the Leguminosae exhibit a similar behavior. For instance, Glc/Man-binding lectins from the tribe Vicieae are best inhibited by biantennary glycans of the *N*-acetyllactosaminic type bearing an α1,6-linked Fuc on the first GlcNAc residue (Debray et al., 1981; Debray and Rougé, 1984). These results were supported by further crystallographic analyses performed on the *Lathyrus ochrus* lectins LoLI and LoLII complexed to various oligosaccharides or glycoproteins (Bourne et al., 1990, 1992, 1994).

The noncovalent association of monomers into dimers or tetramers with multiple Man-binding sites is apparently a widespread feature that has also been observed for other

sugar-binding proteins (Drickamer, 1995). By virtue of this multivalency, the monocot Man-binding lectins are probably capable of interacting more strongly with either simple or complex Man-containing glycoconjugates. In this context, it is interesting to note that the tetrameric snowdrop lectin is a more potent insecticidal lectin than the trimeric daffodil lectin and the dimeric garlic lectin (Powell et al., 1995).

Received June 13, 1996; accepted September 12, 1996.

LITERATURE CITED

Balzarini J, Neyts J, Schols D, Hosoya M, Van Damme EJM, Peumans WJ, De Clercq E (1992) The mannose-specific plant lectins from *Cymbidium hybrid* and *Epipactis helleborine* and the (N-acetylglucosamine)-specific plant lectin from *Urtica dioica* are potent and selective inhibitors of human immunodeficiency virus and cytomegalovirus replication in vitro. Antiviral Res **18:** 191–207

Balzarini J, Schols D, Neyts J, Van Damme EJM, Peumans WJ, De Clercq E (1991) α-(1–3)- and α-(1–6)-D-mannose-specific plant lectins are markedly inhibitory to human immunodeficiency virus and cytomegalovirus infections in vitro. Antimicrob Agents Chemother **35:** 410–416

Barbieri L, Battelli MG, Stirpe F (1993) Ribosome-inactivating proteins from plants. Biochim Biophys Acta **1154:** 237–282

Bourne Y, Mazurier J, Legrand D, Rougé P, Montreuil J, Spik G, Cambillau C (1994) Interaction of a legume lectin with the human lactotransferrin N2 fragment or with the isolated biantennary glycopeptide: role of the fucose moiety. Structure **2:** 209–219

Bourne Y, Rougé P, Cambillau C (1990) X-ray structure of α(α-Man(1–3)β-Man(1–4)GlcNAc)-lectin complex at 2.1 Å resolution. The role of water in sugar-lectin interaction. J Biol Chem **265:** 18161–18165

Bourne Y, Rougé P, Cambillau C (1992) X-ray structure of a biantennary octasaccharide-lectin complex refined at 2.3 Å resolution. J Biol Chem **267:** 197–203

Dahlgren RMT, Clifford HT, Yeo PF (1985) The Families of the Monocotyledons: Structure, Evolution and Taxonomy. Springer-Verlag, Berlin

Debray H, Decout D, Strecker G, Spik G, Montreuil J (1981) Specificity of twelve lectins towards oligosaccharides and glycopeptides related to N-glycosylproteins. Eur J Biochem **117:** 41–55

Debray H, Rougé P (1984) The fine sugar specificity of the *Lathyrus ochrus* seed lectin and isolectins. FEBS Lett **176:** 120–124

Dill K, Olson JD (1995) Picogram detection levels of asialofetuin via the carbohydrate moieties using the light addressable potentiometric sensor. Glycoconjugate Journal **12:** 660–663

Drickamer K (1995) Multiplicity of lectin-carbohydrate interactions. Nature Struct Biol **2:** 437–439

Gaboriaud C, Bissery V, Benchetrit T, Mornon JP (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Lett **224:** 149–155

Hester G, Kaku H, Goldstein IJ, Wright CS (1995) Structure of mannose-specific snowdrop (*Galanthus nivalis*) lectin is representative of a new plant lectin family. Nature Struct Biol **2:** 472–479

Hilder VA, Powell KS, Gatehouse AMR, Gatehouse JA, Gatehouse LN, Shi Y, Hamilton WDO, Merryweather A, Newell C, Timans JC, Peumans WJ, Van Damme EJM, Boulter D (1995) Expression of snowdrop lectin in transgenic tobacco plants results in added protection against aphids. Transgenic Research **4:** 18–25

Hopp TP, Woods KR (1981) Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci USA **78:** 3824–3828

Janin J (1979) Surface and inside volumes in globular proteins. Nature **277:** 491–492

Kaku H, Van Damme EJM, Peumans WJ, Goldstein IJ (1990) Carbohydrate-binding specificity of the daffodil (*Narcissus pseudonarcissus*) and amaryllis (*Hippeastrum hybr.*) bulb lectins. Arch Biochem Biophys **279:** 298–304

Kaku H, Van Damme EJM, Peumans WJ, Goldstein IJ (1992) New mannose-specific lectins from garlic (*Allium sativum*) and ramsons (*Allium ursinum*) bulbs. Carbohydr Res **229:** 347–353

Karplus PA, Schulz GE (1985) Prediction of chain flexibility in proteins. Naturwissenschaften **72:** 212

Koike T, Titani K, Suzuki M, Beppu H, Kuzuya H, Maruta K, Shimpo K, Fujita K (1995) The complete amino acid sequence of a mannose-binding lectin from "Kidachi aloe" (*Aloe arborescens* var. Natalensis Berger). Biochem Biophys Res Commun **214:** 163–170

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol **157:** 105–132

Lemesle-Varloot L, Henrissat B, Gaboriaud C, Bissery V, Morgat A, Mornon JP (1990) Hydrophobic cluster analysis: procedure to derive structural and functional information from 2-D-representation of protein sequences. Biochimie **72:** 555–574

Maddison WP, Maddison DR (1992) MacClade: Analysis of Phylogeny and Character Evolution, version 3.0. Sinauer Associates, Sunderland, MA

Parker JMR, Guo D, Hodges RS (1986) New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry **25:** 5425–5432

Peumans WJ, Van Damme EJM (1995) Lectins as plant defense proteins. Plant Physiol **109:** 347–352

Powell KS, Gatehouse AMR, Hilder VA, Van Damme EJM, Peumans WJ, Boonjawat J, Horsham K, Gatehouse JA (1995) Different antimetabolic effects of related lectins towards nymphal stages of *Nilaparvata lugens*. Entomol Exp Appl **75:** 61–65

Pusztai A, Grant G, Spencer RJ, Duguid TJ, Brown DS, Ewen SWB, Peumans WJ, Van Damme EJM, Bardocz S (1993) Kidney bean lectin-induced *E. coli* overgrowth in the small intestine is blocked by GNA, a mannose-specific lectin. J Appl Bacteriol **75:** 360–368

Raikhel NV, Broekaert WF (1993) Structure and function of chitin-binding proteins. Annu Rev Plant Physiol Plant Mol Biol **44:** 591–615

Saito K, Komae A, Kakuta M, Van Damme EJM, Peumans WJ, Goldstein IJ, Misaki A (1993) The alpha-mannosyl-binding lectin from leaves of the orchid twayblade (*Listera ovata*): application to separation of α-D-mannans from α-D-glucans. Eur J Biochem **217:** 677–681

Sharon N, Lis H (1990) Legume lectins—a large family of homologous proteins. FASEB J **4:** 3198–3208

Shibuya N, Goldstein IJ, Van Damme EJM, Peumans WJ (1988) Binding properties of a mannose-specific lectin from the snowdrop (*Galanthus nivalis*) bulb. J Biol Chem **263:** 728–734

Shinohara Y, Kim F, Shimizu M, Goto M, Tosu M, Hasegawa Y (1994) Kinetic measurement of the interaction between an oligosaccharide and lectins by a biosensor based on surface plasmon resonance. Eur J Biochem **223:** 189–194

Smeets K, Van Damme EJM, Peumans WJ (1994) Comparative study of the post-translational processing of the mannose-binding lectins in the bulbs of garlic (*Allium sativum*) and ramsons (*Allium ursinum*). Glycoconjugate Journal **11:** 309–320

Sturm A, Bergwerff AA, Vliegenthart FG (1992) ^{1}H-NMR structural determination of the N-linked carbohydrate chains on glycopeptides obtained from the bean lectin phytohemagglutinin. Eur J Biochem **204:** 313–316

Thornton JM, Edwards MS, Taylor WR, Barlow DJ (1986) Location of continuous antigenic determinants in the protruding regions of proteins. EMBO J **5:** 409–413

Van Damme EJM, Allen AK, Peumans WJ (1987) Isolation and characterization of a lectin with exclusive specificity towards

mannose from snowdrop (*Galanthus nivalis*) bulbs. FEBS Lett **215:** 140–144

Van Damme EJM, Allen AK, Peumans WJ (1988) Related mannose-specific lectins from different species of the family Amaryllidaceae. Physiol Plant **73:** 52–57

Van Damme EJM, Balzarini J, Smeets K, Van Leuven F, Peumans WJ (1994a) The monomeric and dimeric mannose binding proteins from the Orchidaceae species *Listera ovata* and *Epipactis helleborine*: sequence homologies and differences in biological activities. Glycoconjugate Journal **11:** 321–332

Van Damme EJM, Barre A, Rougé P, Van Leuven F., Balzarini J, Peumans WJ (1996a) Molecular cloning of the lectin and a lectin related protein from common Solomon's seal (*Polygonatum multiflorum*). Plant Mol Biol **31:** 657–672

Van Damme EJM, Briké F, Winter HC, Van Leuven F, Goldstein IJ, Peumans WJ (1996b) Molecular cloning of two different mannose-binding lectins from tulip bulbs. Eur J Biochem **236:** 419–427

Van Damme EJM, Declercq N, Claessens F, Hemschoote K, Peeters B, Peumans WJ (1991a) Molecular cloning and characterization of multiple isoforms of the snowdrop (*Galanthus nivalis* L.) lectin. Planta **186:** 35–43

Van Damme EJM, Goldstein IJ, Vercammen G, Vuylsteke J, Peumans WJ (1992a) Lectins of members of the Amaryllidaceae are encoded by multigene families which show extensive homologies. Physiol Plant **86:** 245–252

Van Damme EJM, Goossens K, Smeets K, Van Leuven F, Verhaert P, Peumans WJ (1995) The major tuber storage protein of Araceae species is a lectin. Plant Physiol **107:** 1147–1158

Van Damme EJM, Kaku H, Perini F, Goldstein IJ, Peeters B, Yagi F, Decock B, Peumans WJ (1991b) Biosynthesis, primary structure and molecular cloning of snowdrop (*Galanthus nivalis* L.) lectin. Eur J Biochem **202:** 23–30

Van Damme EJM, Peumans WJ (1988) Biosynthesis of the snowdrop (*Galanthus nivalis*) lectin in ripening ovaries. Plant Physiol **86:** 922–926

Van Damme EJM, Smeets K, Balzarini J, Pusztai A, Van Leuven F, Goldstein IJ, Peumans WJ (1993a) Cloning and characterization of the lectin cDNA clones from onion, shallot and leek. Plant Mol Biol **23:** 365–376

Van Damme EJM, Smeets K, Torrekens S, Van Leuven F, Goldstein IJ, Peumans WJ (1992b) The closely related homomeric and heterodimeric mannose-binding lectins from garlic are encoded by one-domain and two-domain lectin genes, respectively. Eur J Biochem **206:** 413–420

Van Damme EJM, Smeets K, Torrekens S, Van Leuven F, Peumans WJ (1993b) The mannose-specific lectins from ramsons (*Allium ursinum* L.) are encoded by three sets of genes. Eur J Biochem **217:** 123–129

Van Damme EJM, Smeets K, Torrekens S, Van Leuven F, Peumans WJ (1994b) Characterization and molecular cloning of the mannose binding lectins from three Orchidaceae species: *Listera ovata, Epipactis helleborine* and *Cymbidium hybrid*. Eur J Biochem **221:** 769–777

Van Damme EJM, Smeets K, Van Leuven F, Peumans WJ (1994c) Molecular cloning of mannose binding lectins from *Clivia miniata*. Plant Mol Biol **24:** 825–830