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

Corinne HOULÈS ASTOUL*, Willy J. PEUMANS†, Els J. M. VAN DAMME†, Annick BARRE*, Yves BOURNE‡ and Pierre ROUGÉ*[1]

*Institut de Pharmacologie et Biologie Structurale, UMR-CNRS 5089, 205 Route de Narbonne, 31077 Toulouse Cedex 4, France, †Laboratory for Phytopathology and Plant Protection, Katholieke Universiteit Leuven, Willem De Croylaan 42, 3001 Leuven, Belgium, and ‡Architecture et Fonction des Macromolécules Biologiques, UMR-CNRS 9039, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France

Mannose-specific lectins with high sequence similarity to jacalin and the *Maclura pomifera* agglutinin have been isolated from species belonging to the families Moraceae, Convolvulaceae, Brassicaceae, Asteraceae, Poaceae and Musaceae. Although these novel mannose-specific lectins are undoubtedly related to the galactose-specific Moraceae lectins there are several important differences. Apart from the obvious differences in specificity, the mannose- and galactose-specific jacalin-related lectins differ in what concerns their biosynthesis and processing, intracellular location and degree of oligomerization of the protomers. Taking into consideration that the mannose-specific lectins are widely distributed in higher plants, whereas their galactose-specific counterparts are confined to a subgroup of the Moraceae sp. one can reasonably assume that the galactose-specific Moraceae lectins are a small-side group of the main family. The major change that took place in the structure of the binding site of the diverging Moraceae lectins concerns a proteolytic cleavage close to the N-terminus of the protomer. To corroborate the impact of this change, the specificity of jacalin was re-investigated using surface plasmon resonance analysis. This approach revealed that in addition to galactose and *N*-acetylgalactosamine, the carbohydrate-binding specificity of jacalin extends to mannose, glucose, *N*-acetylmuramic acid and *N*-acetylneuraminic acid. Owing to this broad carbohydrate-binding specificity, jacalin is capable of recognizing complex glycans from plant pathogens or predators.

Key words: carbohydrate-binding proteins, galactose, mannose, plant proteins, processing.

INTRODUCTION

Evidence has been presented during the last few decades that many flowering plants accumulate carbohydrate-binding proteins in seeds and/or different vegetative tissues [1]. Biochemical, molecular and structural studies revealed that these so-called 'lectins' or 'agglutinins' can be classified into seven distinct families of evolutionary and structurally related proteins, which implies that plants developed only a limited number of carbohydrate-binding motifs [1]. Detailed analyses of the specificity of numerous lectins further indicated that most of these carbohydrate-binding motifs preferentially or even exclusively recognize foreign glycans. Based on these observations, the concept was developed that most plant lectins do not play a specific role in growth and development of the plant, but are involved in the recognition of 'foreign' glycans [1]. This obvious ability to recognize and bind typical animal glycans has been interpreted as a strong argument in favour of a defensive role of plant lectins against herbivorous animals and phytophageous invertebrates [1].

Within most lectin families, the structure of the carbohydrate-binding site(s) is well conserved and accordingly these lectin families exhibit a rather narrow range of specificities. For example, all lectins containing hevein-like domains (e.g. wheat-germ agglutinin) specifically bind to chitin and *N*-acetylglucosamine oligomers [2]. Similarly, all amaranthins (e.g. lectin from *Amaranthus caudatus*, which exhibits a β-trefoil structure) [3] are specifically inhibited by *N*-acetylgalactosamine (GalNAc) and the T-antigen (Gal-β1,3-GalNAc) [4]. The same holds true for the extended family of the monocot mannose-binding lectins, which all have a similar β-prism structure and an exclusive specificity towards mannose [5]. A fairly strict conservation of structure and specificity also applies to the B chain of type 2 ribosome-inactivating proteins. All documented examples of B chains consist of two covalently linked domains, each of which harbours a single site that binds galactose or GalNAc. Only a few type 2 ribosome-inactivating proteins of the genus *Sambucus* possess one site with an exclusive specificity towards NeuAc(α-2,6)Gal/GalNAc-containing glycans [6,7]. In contrast with the above-mentioned families, the legume lectins display a broad range of distinct carbohydrate-binding specificities notwithstanding the fact that they all are dimers or tetramers possessing the canonical β-sandwich structure [8,9]. This apparent discrepancy between a strict conservation of structure and variation in specificity apparently relies on substitutions of a few amino acids involved in sugar-binding activity and variations in length of a particular loop, which profoundly change the structure of the binding site without affecting the overall three-dimensional structure of the protomers [10,11]. Besides the legume lectin family, heterogeneity with respect to carbohydrate-binding specificity was also found in the family of jacalin-related lectins (JRLs). Previously [12,13], JRLs were considered a small homo-

Abbreviations used: BanLec, Banana lectin; Calsepa, *Calystegia sepium* agglutinin; Conarva, *Convolvulus arvensis* agglutinin; diMan1,2, α1-2 mannobiose; HCA, hydrophobic cluster analysis; Heltuba, *Helianthus tuberosus* agglutinin; JRL, jacalin-related lectin; gJRL, galactose-specific JRL; MeMan, α-methyl mannopyranoside; mJRL, mannose-specific JRL; MPA, *Maclura pomifera* agglutinin; Orysata, *Oryza sativa* agglutinin; SBA, soya bean agglutinin; SPR, surface plasmon resonance.

[1] To whom correspondence should be addressed (e-mail Pierre.Rouge@ipbs.fr).

© 2002 Biochemical Society

geneous family of galactose/T-antigen-binding agglutinins with a typical β-prism structure. Besides jacalin from jack fruit (*Artocarpus integrifolia*) [14] similar lectins were identified in a few other *Artocarpus* sp. and in the Osage orange (*Maclura pomifera*) [15]. However, previous studies on the characterization and cloning of lectins from *Calystegia sepium* (Convolvulaceae) [16,17], *Helianthus tuberosus* (Asteraceae) [18], *Oryza sativa* (Poaceae) [19,20], *Musa acuminata* (Musaceae) [21,22], *Castanea crenata* (family Fagaceae) [23], *Parkia platycephala* (family Fabaceae) [24] and *Brassica napus* (family Brassicaceae), plants [25] revealed that the galactose/T-antigen-specific Moraceae lectins represent only a small subgroup of an extended family of JRLs, most of which exhibit an exclusive specificity towards mannose or oligomannosides. At present, the JRLs are divided into two subfamilies grouping galactose- and mannose-specific homologues respectively. The major difference (apart from the specificity) between the galactose-specific JRLs (gJRLs) and mannose-specific JRLs (mJRLs) concerns the structure of the protomers. Jacalin and other galactose-specific homologues are built up of protomers consisting of a heavy (α) and a light (β) polypeptide chain. Both chains are derived from a single precursor from which an N-terminal propeptide and an internal linker tetrapeptide (located between the α- and β-chain) are post-translationally excised [26]. Interestingly, the generation of a glycine residue (at the N-terminus of the α-chain) with a free amino group is essential for the formation of the galactose-binding site and requires the correct excision of the linker peptide [12]. Since the mJRLs consist of 'intact' protomers, it is evident that the glycine residue homologue to the N-terminal glycine of the α-chain of jacalin has no free amino group. To address the question how the processing of JRLs determines the structure–function relationships, the specificity of jacalin was re-investigated in detail and compared with that of its mannose-specific homologues.

MATERIALS AND METHODS

Biochemicals

The sugars diMan1,3 (α1-3 mannobiose), diMan1,6 (α1-6 mannobiose), triMan3,6 (α1-3,α1-6 mannotriose) and pentaMan3,6 (α1-3,α1-6 mannopentaose) were purchased from Dextra Laboratories (Reading, U.K.). Bovine lactotransferrin and human serotransferrin were gifts from Dr H. Debray (UMR-CNRS 111, Villeneuve d'Ascq, France). Other glycoproteins and simple and complex sugars were purchased from Sigma. Human serotransferrin (5 mg) was desialylated in 1 ml of 0.1 M trifluoroacetic acid at 80 °C for 1 h. After freeze-drying, asialo-serotransferrin was dissolved in methanol and freeze-dried. This procedure was repeated three times to eliminate the remaining trifluoroacetic acid.

Isolation of lectins

Soya bean agglutinin (SBA) was isolated from soya bean [*Glycine max* (L.) Merr., cv Hodgson] flour by affinity chromatography as described in [43]. Arcelin-1 was purified from kidney bean (*Phaseolus vulgaris* L. cv RAZ-2) seeds, as described in [44]. The purity of the lectin and arcelin preparations was confirmed by SDS/PAGE in the presence of 2-mercaptoethanol.

The *Calystegia sepium* agglutinin (Calsepa), *Convolvulus arvensis* agglutinin (Conarva) and *Helianthus tuberosus* agglutinin (Heltuba) were isolated as described previously [17,18]. Banana lectin (BanLec) was isolated from ripe fruits, in accordance with the procedure described in [22].

Jacalin (used for the specificity studies) was isolated from jack fruit seeds by affinity chromatography on immobilized galactose. Briefly, 20 g of dry seeds were soaked overnight in distilled water at 2 °C and homogenized in a Waring blender in 200 ml of 0.1 M Tris/HCl (pH 7.5), containing 0.2 M NaCl. The homogenate was centrifuged at 9000 ***g*** for 15 min and the resulting supernatant filtered through filter paper (Whatman 3MM). The cleared extract was applied on to a column (2.5 cm × 10 cm; approx. 50 ml of bed volume) of galactose–Sepharose 4B equilibrated with 0.1 M PBS (pH 7.5). After loading the extract, the column was washed with the same buffer until the A_{280} decreased below 0.01. The bound lectin was eluted with a solution of 0.1 M galactose in PBS, dialysed against 0.1 M PBS (pH 7.5) and subjected to a second round of affinity chromatography on galactose–Sepharose 4B. The resulting lectin was dialysed against the appropriate buffer and stored at −20 °C until use. Under the conditions used for affinity chromatography, the mannose-binding artocarpin (jack fruit lectin, also known as KM+) does not bind to galactose–Sepharose 4B (but can be recovered from the flow through). The second affinity chromatography on galactose–Sepharose 4B was included to remove artocarpin that was possibly physically trapped during the first round of affinity chromatography.

The binding of jacalin to mannose–Sepharose 4B was studied in separate experiments. Purified jacalin (100 mg), dissolved in 50 ml of 0.1 M PBS (pH 7.5), was applied on to a column (2.5 cm × 10 cm; approx. 50 ml bed volume) of mannose–Sepharose 4B equilibrated with 0.1 M PBS (pH 7.5). After loading, the column was washed with 100 ml of the same buffer and the lectin eluted with either 0.1 M galactose or 0.1 M mannose in PBS.

Surface plasmon resonance (SPR) analysis

SPR analysis was used to study the specific interaction of BanLec, Calsepa, Conarva, Heltuba and jacalin with various glycoproteins (arcelin-1, orosomucoid or α1-acid glycoprotein, ovomucoid, SBA, native and desialylated human serotransferrin, native and desialylated calf fetuin). Analyses were performed using a BIAcore 1000 (Pharmacia Biosensor AB) biosensor system.

Sensor chips (CM 5) and all the chemicals required for the activation of the carboxymethylated dextran and the immobilization of glycoproteins (100 mM *N*-hydroxysuccinimide, 400 mM *N*-ethyl-*N′*-(3-dimethylaminopropyl) carbodi-imide hydrochloride and 1 M ethanolamine hydrochloride adjusted to pH 8.5 with NaOH) were obtained from Pharmacia Biosensor AB. Hepes-buffered saline [HBS: 10 mM Hepes (pH 7.4)/150 mM NaCl/3 mM EDTA], containing 0.05% BIAcore surfactant P20 used for the biosensor measurements, was from Pharmacia Biosensor AB. For immobilization, glycoproteins were dissolved in 5 mM sodium acetate (pH 4.0) at a final concentration of 1 mg $\cdot$ ml^{-1}. According to the change in SPR response as a result of the immobilization of the glycoproteins on the carboxymethylated dextran layer covering the sensor chip, the surface concentration of the immobilized glycoproteins was approx. 10 ng $\cdot$ mm^{-2}.

All operations were performed at 25 °C. The lectins, dissolved in HBS at concentrations ranging from 25 to 100 μg $\cdot$ ml^{-1}, were injected for 5 min on to the glycoprotein-bound surface of the sensor chip at a flow rate of 5 μl $\cdot$ min^{-1}. The change of the SPR response [expressed as resonance units (RU)], was monitored for 9.5 min. Inhibition was studied by injecting 5–25 mM solutions of monosaccharides or oligosaccharides in HBS at the beginning of the dissociation phase for 5 min at a flow rate of

© 2002 Biochemical Society

5 μl · min^{-1}, and the change of the SPR response (RU) was monitored for 9.5 min. The extent of inhibition was determined by measuring the percentage of lectin remaining immobilized on the glycoprotein-bound surface.

Molecular modelling and docking

The amino acid sequence alignments were performed with CLUSTAL-X [45]. Hydrophobic cluster analysis (HCA) [46,47] was performed to delineate the structurally conserved regions along the amino acid sequences of JRLs, using Heltuba as a model. HCA plots were generated using the program HCA-Plot2 (Doriane, Paris, France).

Homology modelling of BanLec, Calsepa and Conarva was performed on a Silicon Graphics O2 R10000 workstation, using the programs InsightII, Homology and Discover (Accelrys, San Diego, CA, U.S.A.). The atomic co-ordinates of Heltuba (RCSB Protein Data Bank code 1c3m) were used to build the three-dimensional model of the JRLs. Steric conflicts resulting from the replacement or the deletion of some residues were corrected during the model building procedure using the rotamer library [48] and the search algorithm implemented in the Homology program [49] to maintain proper side-chain orientation. Energy minimization and relaxation of the loop regions were performed by several cycles of steepest descent and conjugate gradient using the consistent valence forcefield of Discover. PROCHECK [50] was used to check the stereochemical assessment of the three-dimensional models. The program TurboFrodo (Bio-Graphics, Marseille, France) was run on the workstation to draw the Ramachandran plots and superimpose the models. Figures were generated using Molscript [51], Bobscript [52] and Raster3D [53].

For docking of mannose into the monosaccharide-binding site of the different JRLs, the lowest apparent binding energy (E_{bind} expressed in kcal · mol^{-1}) compatible with the hydrogen bonds {considering van der Waals interactions and strong [2.5 Å (1 Å $= 10^{-10}$ m) $<$ dist(D–A) $<$ 3.1 Å and 120° $<$ ang(D–H–A)] and weak [2.5 Å $<$ dist(D–A) $<$ 3.5 Å and 105° $<$ ang(D–H–A) $<$ 120°] hydrogen bonds with D, donor; A, acceptor; and H, hydrogen} found in Heltuba complexed with mannose [30] was calculated with a consistent valence forcefield and used to anchor the pyranose ring of mannose into the binding sites. Orientation of the pyranose ring of mannose as it occurs in the Heltuba–mannose complex was used as a starting position for the docking experiments.

RESULTS AND DISCUSSION

The present study deals with a re-investigation of the specificity of jacalin and the corroboration of the structural basis for the promiscuous specificity of an hitherto presumed galactose/T-antigen-binding protein. The principal aim is to demonstrate how a simple proteolytic cleavage of the lectin polypeptide thoroughly changes the specificity of a carbohydrate-binding motif without altering the overall fold and three-dimensional structure of the protomer.

mJRLs and gJRLs

Biochemical and molecular biological studies provided ample evidence for the occurrence of a widespread family of mannose-binding lectins sharing sequence similarity and structural similarity with the classic galactose/T-antigen-specific jacalin. According to a recently proposed model [27] mJRLs are probably ubiquitous among flowering plants and evolved from a common ancestor, which was already present in early lower plants. All currently known mJRLs are built up of similar domains consisting of an intact (i.e. uncleaved) polypeptide of approx. 150 amino acid residues. In contrast, jacalin and all other galactose-specific homologues are composed of 'cleaved' protomers consisting of a short β-chain and a long α-chain, which are both derived from a single precursor and remain non-covalently associated. It has been proposed that the vacuolar gJRLs evolved from a cytoplasmic mannose-specific homologue through the insertion of vacuolar-targeting sequences [27]. Since this evolutionary event has taken place after the Moraceae diverged from other flowering plants, the occurrence of the gJRLs is confined to a small taxonomic group.

The insertion of vacuolar-targeting sequences and the concomitant post-translational processing profoundly affect the structure of the binding site of the gJRLs, because the sugar-binding pocket is not located on a single peptide (as in the mJRLs) but is spread over two separate chains. One can reasonably expect, therefore, that the specificity of the gJRLs is thoroughly changed as compared with that of the mJRLs. To corroborate the consequences of the cleavage of the lectin domain on the structure and specificity of the binding site, the specificity of jacalin was re-investigated and the results revisited in view of the structural similarities/differences between mJRLs and gJRLs.

The specificity of jacalin revisited: the binding site not only accommodates galactose/T-antigen but also mannose

In spite of the early observation that the agglutination activity of jacalin and *M. pomifera* agglutinin (MPA) was (besides galactose) also inhibited by mannose [28,29], both lectins are classically considered as lectins with an exclusive specificity towards galactose/T-antigen [12,13]. This canonical specificity of jacalin/MPA has never been questioned until structural analyses unambiguously demonstrated that the binding site of some newly discovered mannose-specific homologues are structurally very similar to that of jacalin [30,31]. In a first attempt to check the possible binding activity of jacalin towards mannose, the purified lectin was applied on to a column of immobilized mannose. Thereby it was observed that jacalin was quantitatively retained on immobilized mannose (results not shown). Moreover, jacalin bound to galactose–Sepharose 4B could be eluted with a 0.1 M solution of mannose. Similarly, jacalin bound to mannose–Sepharose 4B was completely desorbed with a 0.1 M solution of galactose. Since these simple-affinity chromatography experiments indicated that jacalin has a broader specificity than is generally believed, the carbohydrate-binding properties of jacalin were re-investigated by SPR analysis of the interaction between the lectin and immobilized arcelin in the presence of various monosaccharides and oligosaccharides.

As shown in Figure 1, the inhibition experiments unambiguously demonstrated that besides galactose also α-methyl mannopyranoside (MeMan), mannose, glucose, *N*-acetylneuraminic acid and *N*-acetylmuramic acid, but not L-fucose competed for the interaction of jacalin with arcelin.

These results leave no doubt that the monosaccharide-binding specificity of jacalin is not restricted to galactose or GalNAc, but extends to its epimers, e.g. glucose and mannose. Most probably, the same holds true for MPA, which exhibits specificity virtually identical with that of jacalin [13,29]. To confirm these unexpected results, docking experiments were performed to check whether the carbohydrate-binding site of jacalin can accommodate mannose or glucose. Thereby, it became evident that jacalin can,

© 2002 Biochemical Society

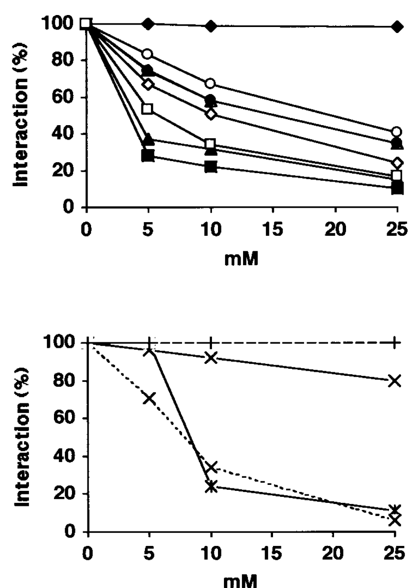

Figure 1 SPR analysis of the specificity of jacalin towards galactose, mannose and oligomannosides

Graphs show the interaction of jacalin with immobilized arcelin in the presence of galactose (◆), mannose (○), glucose (x), MeMan (■), L-fucose (---+---), *N*-acetylneuraminic acid (---x---), *N*-acetylmuramic acid (⁎), diMan1,2 (△), diMan1,3 (●), diMan1,6 (◇), triMan3,6 (▲) and pentaMan3,6 (□) added during the dissociation phase at increasing concentrations ranging from 5 to 25 mM. Results are expressed as the percentage of the lectins retained on immobilized arcelin. All values are the means of triplicate experiments.

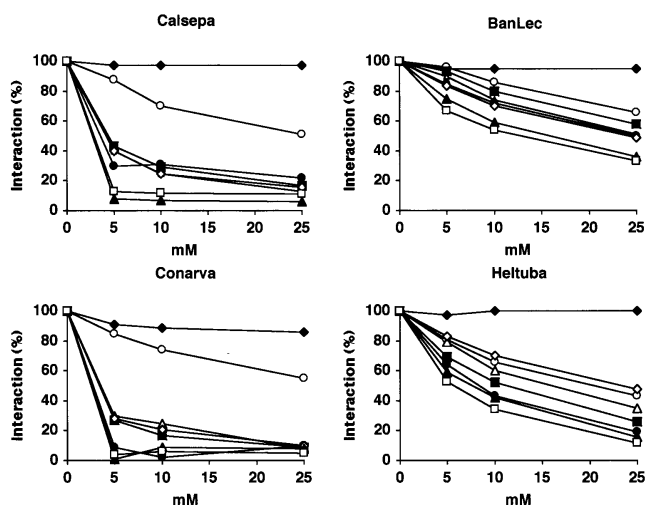

Figure 2 SPR analysis of the specificity of jacalin and JRL towards mannose and oligomannosides

The interaction of JRLs and jacalin with immobilized arcelin in the presence of mannose (○), MeMan (■), galactose (◆), diMan1,2 (△), diMan1,3 (●), diMan1,6 (◇), triMan3,6 (▲) and pentaMan3,6 (□) added during the dissociation phase at increasing concentrations ranging from 5 to 25 mM. Results are expressed as the percentage of the lectins retained on immobilized arcelin. All values are the means of triplicate experiments.

indeed, bind mannose and glucose, even though the network of hydrogen bonds connecting these sugars to the amino acid residues of the site is slightly different and seems weaker than that formed with galactose [12].

mJRLs exhibit marked differences in fine specificity

The unexpected results obtained with jacalin also raised some questions about the presumed exclusive specificity of the mJRLs towards mannose and possible differences in fine specificity between different mJRLs. Therefore a detailed study was made on the interaction of four mJRLs (BanLec, Calsepa, Conarva and Heltuba) with complex glycan chains of various glycoproteins using SPR measurements. The immobilized glycoproteins used in these experiments contain predominantly glycans of the high-mannose type (arcelin-1, SBA, ovalbumin and ovomucoid) and complex glycans of the *N*-acetyllactosaminic type (orosomucoid and human serotransferrin) [32]. All these glycans contain a more or less exposed trimannoside core Manα1,6(Manα1,3) Man, which was shown to interact specifically with the extended carbohydrate-binding site of many lectins such as concanavalin A [33]. In addition, fetuin contains a mixture of N- and O-linked glycans containing sialylated terminal galactose residues [34,35].

The results of the SPR measurements are summarized in Figure 2. Basically, the mJRLs can be subdivided into two distinct groups according to their reactivity towards glycoproteins. Conarva and Calsepa reacted poorly with glycoproteins containing exclusively (SBA) or predominantly (arcelin) high-mannose-type glycan chains, whereas they readily interacted with glycoproteins containing a mixture of high-mannose and complex type glycans (in casu sialylated and desialylated human serotransferrin, fetuin and asialofetuin and orosomucoid). In contrast, BanLec and Heltuba interacted strongly with immobilized arcelin and SBA, but exhibited a very weak affinity towards other glycoproteins. Accordingly, the estimated K_D values of the lectins towards the high-mannose-type glycans of arcelin measured by SPR were of a different order of magnitude for Heltuba (4.0×10^{-8} M), and BanLec (5.4×10^{-9} M) compared with Conarva (1.1×10^{-6} M) and Calsepa (1.3×10^{-6} M). The apparent lower reactivity of the lectins towards SBA is probably due to the lower (as compared with arcelin) content of high-mannose glycans [36] combined with a reduced accessibility as a result of its immobilization on to the CM-dextran surface of the sensor chip. It is worth noting in this respect that the protrusion of the four Man9-GlcNAc N-glycans linked to Asn^{75} residues of SBA in a single plane from the dimer–dimer interface of native (tetrameric) SBA [37] may cause considerable steric hindrance upon immobilization. The N-glycans of the arcelin-1 dimer, which protrude in different orientations in the solvent [38], can interact better with the lectins after immobilization.

The fairly strong reaction of Calsepa and Conarva with sialylated and desialylated human serotransferrin and calf fetuin most probably relies on the presence of an exposed trimannoside core in the *N*-acetyl-lactosaminic-type glycans of these glycoproteins. It is worth mentioning in this context that also some mannose-binding legume lectins readily recognize this core structure [33,39].

The marked affinity of the mJRLs towards mannose and mannose-containing oligosaccharides was further corroborated by SPR using inhibition experiments of the lectin–arcelin interaction by mannose, mannose derivatives and a series of α-mannosides of increasing size. As shown in Figure 3, galactose, even at concentrations as high as 25 mM, had no inhibitory effect on the binding of BanLec, Calsepa, Conarva and Heltuba to arcelin, whereas mannose, glucose and their derivatives readily inhibited the lectin–arcelin interaction at concentrations ranging between 5 and 25 mM.

Apart from a few discrepancies (e.g. Heltuba and BanLec which are equally well inhibited by α-methyl glucopyranoside and mannose) the relative inhibitory potency of simple sugars is

© 2002 Biochemical Society

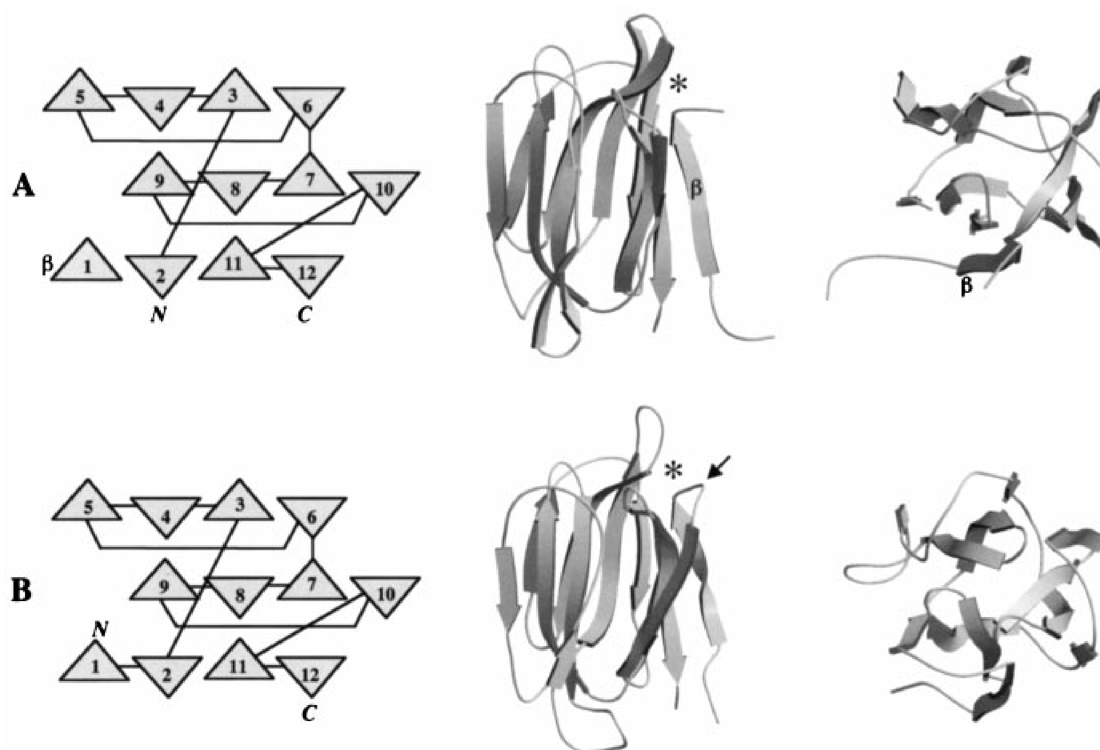

Figure 3 Comparison of the three-dimensional structure of jacalin (A) and BanLec (B)

Figures on the left side indicate the arrangement of the 12 strands of β-sheet (grey triangles) forming the three four-stranded bundles of the lectin protomers. A free strand of β-sheet corresponding to the β-chain (β) completes one of the four-stranded bundles of jacalin, whereas all the strands of other mannose-specific lectins are located on a single polypeptide chain. Strands of β-sheet are numbered according to their occurrence along the polypeptide chain (starting from the N-terminus) and that corresponding to the light (β) chain is numbered 1. The three-dimensional models shown in the middle and on the right correspond to the side and top views of the monomers respectively. Strands of β-sheet are indicated by arrows. The asterisks and arrow indicate the location of the carbohydrate-binding sites and the position of the linker peptide connecting the first strand of β-sheet to the rest of the polypeptide chain of BanLec respectively. This linker peptide is lacking in jacalin and MPA. Cartoons were generated using Molscript [51], Bobscript [52] and Raster3D [53].

similar for all four lectins, namely, MeMan > α-methyl glucopyranoside > mannose > glucose > galactose. Inhibition experiments with oligomannosides showed that triMan3,6 and pentaMan3,6 are the best inhibitors for all the lectins assayed, followed by different dimannosides, MeMan and mannose. There are, however, some discrepancies concerning α1-2 mannobiose (diMan1,2), because this disaccharide is a better inhibitor of Calsepa than diMan1,3 or diMan1,6. It should be emphasized that diMan1,2 linkages, which are the most exposed regions of the high-mannose-type glycans, interact very strongly with some mJRLs. The results from these inhibition assays clearly demonstrate that the carbohydrate-binding site of the mJRLs can accommodate more than a single sugar unit. Moreover, since pentaMan3,6 is not a more potent inhibitor than triMan3,6, one can reasonably assume that the binding site accommodates not more than two or three mannose units.

Structural basis for the differences in specificity between gJRLs and mJRLs: mannose- and galactose-binding JRLs share the same overall fold and three-dimensional structure, but possess structurally different sugar-binding sites

To relate the obvious differences in specificity to differences in structural features a comparison was made of the overall fold and three-dimensional structure, and the structure of the binding sites of gJRLs and mJRLs. At present, the three-dimensional structures of two gJRLs, namely jacalin [12] and MPA [13], have been resolved by X-ray crystallography. Both lectin monomers exhibit a very similar β-prism fold, which consists of 12 strands of antiparallel β-sheet connected by loops and clustered into three subdomains corresponding to three bundles of four strands, each oriented parallel to the axis of the prism (Figure 3A). The four β-strands of each subdomain constitute a Greek key motif. This barrel-shaped β-prism structure definitely differs from that of the monocot mannose-binding lectins [40], where the bundles of β-sheet are oriented perpendicularly to the axis of the prism and adopt a flattened star-shaped β-prism structure. The β-prism architecture built by 11 of the 12 strands of β-sheet, of jacalin and MPA, belong to the α-chain and form two complete four-stranded bundles and one incomplete three-stranded bundle respectively. The 12th strand of β-sheet (called β1), which completes the third four-stranded bundle, is located in the β-chain. Each protomer contains a single carbohydrate-binding site located at one end of the barrel-shaped β-prism structure. The four amino acid residues forming this site are the N-terminal glycine of the α-chain and three other residues located at the C-terminus of the α-chain, namely Tyr^{122}, Trp^{123} and Asp^{125}. Hitherto, the three-dimensional structure of two mJRLs, namely Heltuba [30] and artocarpin [41], has been resolved. Heltuba exhibits virtually the same overall fold as jacalin and MPA (Figure 3B). However, since the protomers of Heltuba are not proteolytically processed, the first strand of β-sheet remains

© 2002 Biochemical Society

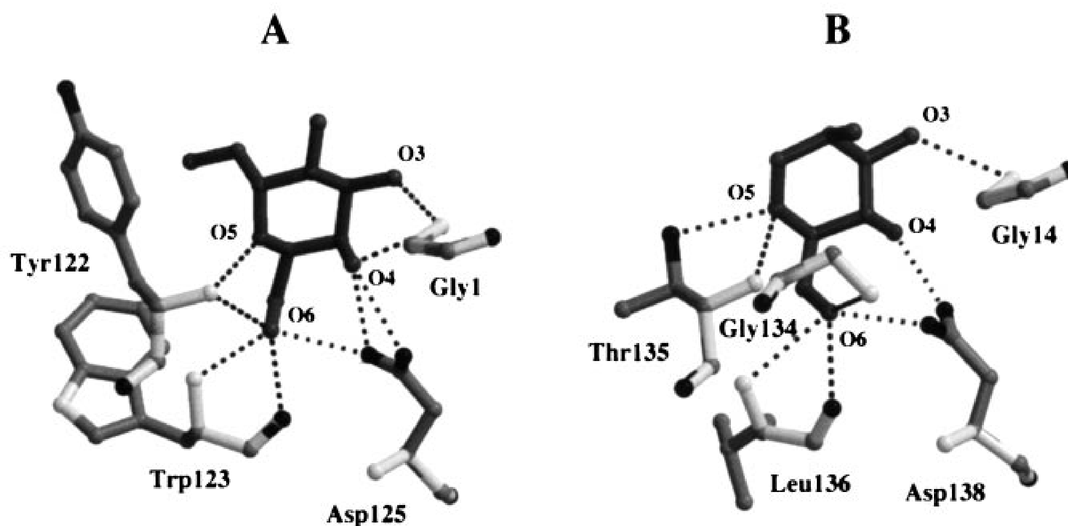

Figure 4 Comparison of the carbohydrate-binding sites of jacalin (A) and Orysata (B)

Hydrogen bonds forming the network connecting galactose (boldface) to the binding site of jacalin and mannose (boldface) to the binding site of Orysata are indicated by dotted lines. Gly^{1} residue of jacalin, which possesses a free -NH_2 group, corresponds to the N-terminus of the α-chain. Its homologue in Orysata corresponds to Gly^{14} residue, which is located on the same polypeptide chain as the other residues forming the binding site of this lectin. Cartoons were generated using Molscript [51], Bobscript [52] and Raster3D [53].

covalently linked to the rest of the polypeptide chain, whereas in jacalin and MPA the β-chain carrying the β1-sheet is cleaved from the α-chain. In the absence of proteolytic cleavage, this first strand of β-sheet remains connected to the core of the polypeptide chain by a rather extended glycine-rich loop, which also forms the edge of the carbohydrate-binding site. The same conclusions can be drawn from the three-dimensional structure of artocarpin [40] and the three-dimensional models built for the mJRLs Conarva, Calsepa, BanLec, *Oryza sativa* agglutinin (Orysata) [20] and artocarpin (KM+) [42]. It is evident, therefore, that the differences in specificity between the gJRLs and mJRLs cannot be attributed to differences in the overall fold and three-dimensional structure and, hence, must be determined by structural differences between the binding sites. Detailed information about the binding sites is only available for jacalin, MPA, Heltuba and artocarpin. A comparison of these structures revealed that the mannose-binding site of Heltuba (in crystals soaked with di-, tri- and penta-mannosides) is located at exactly the same position as the T-antigen-binding site of jacalin and MPA [12,13,30,40]. The mannose-binding site of Heltuba comprises three exposed loops, which are located at the top of the β-prism fold and belong to Greek key motifs 1 and 3 (Figure 3B). Residues Gly^{135}-Asp^{136}-Val^{137} and Asp^{139}, which are located on the surface-exposed β11–β12 loops, create a network of hydrogen bonds with O6, O5 and O4 of mannose. The Gly^{18} residue, located within the β1–β2 loop, forms an additional hydrogen bond with O3 atom and van der Waals interactions with the equatorial O4 conformation of mannose. In contrast with the mannose-binding motifs of the monocot mannose-binding lectins, Heltuba forms no direct hydrogen bond with the axial O2 atom of mannose [8,9]. Interestingly, Met^{92} residue of the β7–β8 loop stacks to the B-face of the mannose pyranose ring. In addition, the side chain of Gly^{135} establishes van der Waals interactions with the O2 atoms of mannose.

Since all mJRLs exhibit specificity towards mannose and do not recognize galactose, one can reasonably assume that their binding sites resemble that of Heltuba. Molecular modelling and docking experiments with Orysata using mannose as a ligand [20] confirmed that mannose is anchored to the carbohydrate-binding site through a network of nine hydrogen bonds (Figure 4B). The hydrogen bonds connecting O4, O5 and O6 of mannose to Thr^{135}, Leu^{136} and Asp^{138} residues are similar to those found in Heltuba complexed with mannose [30]. This network of hydrogen bonds also resembles that connecting galactose to the corresponding Tyr^{122}, Trp^{123} and Asp^{125} residues of jacalin (Figure 4A) [12]. However, as compared with jacalin, additional hydrogen bonds occur between O5 and Oγ1 of Thr^{135} and between O6 and N of Gly^{134} of Orysata. The main difference between Orysata and jacalin concerns the hydrogen bonds connecting O3 and O4 of the pyranose ring to Gly^{14} (Gly^{1} in jacalin) and Asp^{138} (Asp^{125} in jacalin) residues. A single hydrogen bond occurs, indeed, between Gly^{14} of Orysata and O3 of mannose, whereas two bonds connect Gly^{1} of jacalin to O3 and O4 of galactose. Similarly, a single hydrogen bond connects O4 of mannose to Oδ1 of Asp^{138} of Orysata, whereas two hydrogen bonds link O4 of galactose to Oδ1 and Oδ2 of Asp^{125} in jacalin. Docking experiments also suggest that replacing mannose by glucose has little effect on this network of hydrogen bonds, which is in good agreement with the previously reported interaction of glucose with Orysata in hapten inhibition [20].

The present analysis clearly illustrates that the most important structural differences between the mJRLs and gJRLs concern the size and conformation of the carbohydrate-binding site. As a result of the proteolytic cleavage of the protomer, the size of the carbohydrate-binding site of jacalin and MPA is more extended than that of their mannose-specific homologues with intact protomers. This increased size explains the apparent lack of monosaccharide-binding specificity of the carbohydrate-binding sites of jacalin and MPA. However, other structural features of the carbohydrate-binding site of jacalin, e.g. the length of the loop interconnecting β-sheets, β7 and β8 may also contribute to the unusual broad monosaccharide-binding specificity of gJRLs.

© 2002 Biochemical Society

A closer examination of the structure of the carbohydrate-binding sites reveals several interesting features. Apart from the Gly and Asp residues, which bind to O6 of mannose or glucose and are conserved in all lectins, the two other residues hydrogen-bonded to O6 (i.e. the amino acid residues corresponding to Tyr^{122} and Trp^{123} of jacalin) are highly variable. However, these changes have no apparent effect on the carbohydrate-binding specificity of the site because both residues bind to O6 by their —N=H and C=O backbone groups respectively, irrespective of the conformation and hindrance of their side chains. Nevertheless, some amino acid changes create additional hydrogen bonds. More subtle conformational changes in the orientation of the side chain of e.g. His^{91} in Heltuba can explain the preference of this lectin for the dimannosidic α1,2-linkage over the α1,3- or α1,6-linkages, as deduced from SPR measurements [18].

Kidney bean seeds (*Phaseolus vulgaris* L. cv. RAZ-2) used for the isolation and purification of arcelin-1 were a gift of Dr C. Cardona (CIAT, Cali, Columbia). Dr G. Cornaire is gratefully acknowledged for the isolation and purification of arcelin-1. This work was supported by CNRS and the Conseil Régional de Midi-Pyrénées (C.H.A., A.B. and P.R.), and in part by grants from the Research Fund K.U. Leuven (OT/98/17) and the Fund for Scientific Research-Flanders (FWO grant G.0113.01).

REFERENCES

1 Van Damme, E. J. M., Peumans, W. J., Barre, A. and Rougé, P. (1998) 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
2 Goldstein, I. J. and Poretz, R. D. (1986) Isolation, physicochemical characterization, and carbohydrate-binding specificity of lectins. In The Lectins, Properties, Functions, and Applications in Biology and Medicine (Liener, I. E., Sharon, N. and Goldstein, I. J., eds.), pp. 33–247, Academic Press, Orlando, FL
3 Transue, T. R., Smith, A. K., Mo, H., Goldstein, I. J. and Saper, M. A. (1997) Structure of benzyl T-antigen disaccharide bound to *Amaranthus caudatus* lectin. Nat. Struct. Biol. **10**, 779–783
4 Rinderle, S. J., Goldstein, I. J., Matta, K. L. and Ratcliffe, R. M. (1989) Isolation and characterization of amaranthin, a lectin present in the seeds of *Amaranthus caudatus*, that recognizes the T- (or cryptic T)-antigen. J. Biol. Chem. **264**, 16123–16131
5 Barre, A., Van Damme, E. J. M., Peumans, W. J. and Rougé, P. (1996) Structure–function relationship of monocot mannose-binding lectins. Plant Physiol. **112**, 1531–1540
6 Van Damme, E. J. M., Barre, A., Rougé, P., Van Leuven, F. and Peumans, W. J. (1996) The NeuAc(α-2,6)-Gal/GalNAc binding lectin from elderberry (*Sambucus nigra*) bark is a type 2 ribosome inactivating protein with an unusual specificity and structure. Eur. J. Biochem. **235**, 128–137
7 Van Damme, E. J. M., Roy, S., Barre, A., Citores, L., Mostafapous, K., Rougé, P., Van Leuven, F., Girbes, T., Goldstein, I. J. and Peumans, W. J. (1997) Elderberry (*Sambucus nigra*) bark contains two structurally different Neu5Ac(α2,6)Gal/GalNAc-binding type 2 ribosome-inactivating proteins. Eur. J. Biochem. **245**, 648–655
8 Rini, J. (1995) Lectin structure. Annu. Rev. Biophys. Biomol. Struct. **24**, 551–577
9 Loris, R., Hamelryck, T., Bouckaert, J. and Wyns, L. (1998) Legume lectin structure. Biochim. Biophys. Acta **1383**, 9–36
10 Young, N. M. and Oomen, R. P. (1992) Analysis of sequence variation among legume lectins. A ring of hypervariable residues forms the perimeter of the carbohydrate-binding site. J. Mol. Biol. **228**, 924–934
11 Sharma, V. and Surolia, A. (1997) Analyses of carbohydrate recognition by legume lectins: size of the combining site loops and their primary specificity. J. Mol. Biol. **267**, 433–445
12 Sankaranarayanan, R., Sekar, K., Banerjee, R., Sharma, V., Surolia, A. and Vijayan, M. (1996) A novel mode of carbohydrate recognition in jacalin, a *Moraceae* plant lectin with a β-prism fold. Nat. Struct. Biol. **3**, 596–603
13 Lee, X., Thompson, A., Zhang, Z., Ton-That, H., Biesterfeldt, J., Ogata, C., Xu, L., Johnston, A. Z. and Young, N. M. (1998) Structure of the complex of *Maclura pomifera* agglutinin and the T-antigen disaccharide, Galβ1,3GalNAc. J. Biol. Chem. **273**, 6312–6318
14 Moreira, R. A. and Ainouz, I. L. (1981) Lectins from seeds of jack fruit (*Artocarpus integrifolia* L.): isolation and purification of two isolectins from the albumin fraction. Biol. Plant (Praha) **23**, 186–192
15 Bausch, J. N. and Poretz, R. D. (1977) Purification and properties of the hemagglutinin from *Maclura pomifera* seeds. Biochemistry **16**, 5790–5794
16 Van Damme, E. J. M., Barre, A., Verhaert, P., Rougé, P. and Peumans, W. J. (1996) Molecular cloning of the mitogenic mannose/maltose-specific rhizome lectin from *Calystegia sepium*. FEBS Lett. **397**, 352–356
17 Peumans, W. J., Winter, H. C., Bemer, V., Van Leuven, F., Goldstein, I. J., Truffa-Bachi, P. and Van Damme, E. J. M. (1997) Isolation of a novel plant lectin with an unusual specificity from *Calystegia sepium*. Glycoconj. J. **14**, 259–265
18 Van Damme, E. J. M., Barre, A., Mazard, A.-M., Verhaert, P., Horman, A., Debray, H., Rougé, P. and Peumans, W. J. (1999) Characterization and molecular cloning of the lectin from *Helianthus tuberosus*. Eur. J. Biochem. **259**, 135–142
19 Claes, B., Dekeyser, R., Villarroel, R., Van den Bulcke, M., Bauw, G., Van Montagu, M. and Caplan, A. (1990) Characterization of a rice gene showing organ-specific expression in response to salt stress and drought. Plant Cell **2**, 19–27
20 Zhang, W., Peumans, W. J., Barre, A., Houles Astoul, C., Rovira, P., Rougé, P., Proost, P., Truffa-Bachi, P., Jalali, H. A. and Van Damme, E. J. M. (2000) Isolation and characterization of a jacalin-related mannose-binding lectin from salt-stressed rice (*Oryza sativa*) plants. Planta **210**, 970–978
21 Clendennen, S. K. and May, G. D. (1997) Differential gene expression in ripening banana fruit. Plant Physiol. **115**, 463–469
22 Peumans, W. J., Zhang, W., Barre, A., Houles Astoul, C., Balint-Kurti, P. J., Rovira, P., Rougé, P., May, G. D., Van Leuven, F., Truffa-Bachi, P. and Van Damme, E. J. M. (2000) Fruit-specific lectins from banana and plantain. Planta **211**, 546–554
23 Nomura, K., Nakamura, S., Fujitake, M. and Nakanishi, T. (2000) Complete amino acid sequence of Japanese chestnut agglutinin. Biochem. Biophys. Res. Commun. **276**, 23–28
24 Mann, K., Farias, C. M. S. A., Gallego Del Sol, F., Santos, C. F., Grangeiro, T. B., Nagano, C. S., Cavada, B. S. and Calvete, J. J. (2001) 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**, 4414–4422
25 Geshi, N. and Brandt, A. (1998) Two jasmonate-inducible myrosinase-binding proteins from *Brassica napus* L. seedlings with homology to jacalin. Planta **204**, 295–304
26 Yang, H. and Czapla, T. H. (1993) Isolation and characterization of cDNA clones encoding jacalin isolectins. J. Biol. Chem. **268**, 5905–5910
27 Peumans, W. J., Hause, B. and Van Damme, E. J. M. (2000) The galactose-binding and mannose-binding jacalin-related lectins are located in different sub-cellular compartments. FEBS Lett. **477**, 186–192
28 Sastry, M. V. K., Banerjee, P., Patanjali, S. R., Swamy, M. J., Swarnalatha, G. V. and Surolia, A. (1986) Analysis of the saccharide binding to *Artocarpus integrifolia* lectin reveals specific recognition of T-antigen (β-D-Gal(1,3)D-GalNAc). J. Biol. Chem. **261**, 11726–11733
29 Sarkar, M., Wu, A. M. and Kabat, E. A. (1981) Immunochemical studies on the carbohydrate specificity of *Maclura pomifera* lectin. Arch. Biochem. Biophys. **209**, 204–218
30 Bourne, Y., Zamboni, V., Barre, A., Peumans, W. J., Van Damme, E. J. M. and Rougé, P. (1999) *Helianthus tuberosus* lectin, the crystal structure reveals a widespread scaffold for mannose-binding lectins. Structure Fold. Des. **7**, 1473–1482
31 Bourne, Y., Houlès Astoul, C., Zamboni, V., Peumans, W. J., Menu-Bouaouiche, L., Van Damme, E. J. M., Barre, A. and Rougé, P. (2002) Structural basis for the unusual carbohydrate-binding specificity of jacalin towards galactose and mannose. Biochem. J. **364**, 173–179
32 Montreuil, J. (1984) Spatial conformation of glycans and glycoproteins. Biol. Cell **51**, 115–132
33 Naismith, J. H. and Field, R. A. (1996) Structural basis of trimannoside recognition by concanavalin A. J. Biol. Chem. **271**, 972–976
34 Edge, A. S. B. and Spiro, R. G. (1987) Presence of an *O*-glycosidically linked hexasaccharide in fetuin. J. Biol. Chem. **262**, 16135–16141
35 Cummings, D. A., Hellerqvist, C. G., Harris-Brandts, M., Michnick, S. W., Carver, J. P. and Bendiak, B. (1989) Structures of asparagine-linked oligosaccharides of the glycoprotein fetuin having sialic acid linked to *N*-acetylglucosamine. Biochemistry **28**, 6500–6512
36 Dorland, L., van Halbeek, H., Vliegenthart, F. G., Lis, H. and Sharon, N. (1981) Primary structure of the carbohydrate chain of soybean agglutinin. A reinvestigation by high resolution ^{1}H-NMR spectroscopy. J. Biol. Chem. **256**, 7708–7711
37 Dessen, A., Gupta, D., Sabesan, S., Brewer, C. F. and Sacchettini, J. C. (1995) X-ray crystal structure of the soybean agglutinin cross-linked with a biantennary analog of the blood group I carbohydrate antigen. Biochemistry **34**, 4933–4942
38 Mourey, L., Pédelacq, J. D., Birck, C., Fabre, C., Rougé, P. and Samama, J. P. (1998) Crystal structure of arcelin-1 from *Phaseolus vulgaris* at 1.9 Å resolution. J. Biol. Chem. **273**, 12914–12922
39 Bourne, Y., Mazurier, J., Legrand, D., Rougé, P., Montreuil, J., Spik, G. and 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

© 2002 Biochemical Society

40 Hester, G., Kaku, H., Goldstein, I. J. and Wright, C. S. (1995) Structure of mannose-specific snowdrop (*Galanthus nivalis*) lectin is representative of a new plant lectin family. Nat. Struct. Biol. **2**, 472–479
41 Pratap, J. V., Jeyaprakash, A. A., Rani, P. G., Sekar, K., Surolia, A. and Vijayan, M. (2002) Crystal structures of artocarpin, a Moraceae lectin with mannose specificity, and its complex with methyl-α-D-mannose: implications to the generation of carbohydrate specificity. J. Mol. Biol. **317**, 237–247
42 Rosa, J. C., De Oliveira, P. S. L., Garratt, R., Beltramini, L., Resing, K., Roque-Barreira, M.-C. and Greene, L. J. (1999) KM+, a mannose-binding lectin from *Artocarpus integrifolia*: amino acid sequence, predicted tertiary structure, carbohydrate recognition, and analysis of the β-prism fold. Protein Sci. **8**, 13–24
43 Allen, A. K. and Neuberger, A. (1975) A simple method for the preparation of an affinity adsorbent for soybean agglutinin using galactosamine and CH-Sepharose. FEBS Lett. **50**, 362–364
44 Fabre, C., Causse, H., Mourey, L., Koninkx, J., Rivière, M., Hendriks, H., Puzo, G., Samama, J. P. and Rougé, P. (1998) Characterization and sugar-binding properties of arcelin-1, an insecticidal lectin-like protein isolated from kidney bean (*Phaseolus vulgaris* L. cv RAZ-2) seeds. Biochem. J. **329**, 551–560
45 Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tool. Nucleic Acids Res. **15**, 4876–4882
46 Gaboriaud, C., Bissery, V., Benchetrit, T. and Mornon, J. P. (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Lett. **224**, 149–155
47 Lemesle-Varloot, L., Henrissat, B., Gaboriaud, C., Bissery, V., Morgat, A. and Mornon, J. P. (1990) Hydrophobic cluster analysis: procedure to derive structural and functional information from 2-D-representation of protein sequences. Biochimie **72**, 555–574
48 Ponder, J. W. and Richards, F. M. (1987) Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. J. Mol. Biol. **193**, 775–791
49 Mas, M. T., Smith, K. C., Yarmush, D. L., Aisaka, K. and Fine, R. M. (1992) Modeling the anti-CEA antibody combining site by homology and conformational search. Proteins Struct. Funct. Genet. **14**, 483–498
50 Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemistry of protein structures. J. Appl. Crystallogr. **26**, 283–291
51 Kraulis, P. J. (1991) Molscript: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. **24**, 946–950
52 Esnouf, R. M. (1997) An extensively modified version of Molscript that includes greatly enhanced coloring capabilities. J. Mol. Graphics **15**, 132–134
53 Merritt, E. A. and Bacon, D. J. (1997) Raster3D photorealistic molecular graphics. Methods Enzymol. **277**, 505–524

Received 31 May 2002/17 July 2002; accepted 8 August 2002
Published as BJ Immediate Publication 9 August 2002, DOI 10.1042/BJ20020856

© 2002 Biochemical Society