Research Communication

Effect of Linkage on the Location of Reducing and Nonreducing Sugars Bound to Jacalin

K. V. Abhinav
Kaushal Sharma
Avadhesha Surolia
Mamannamana Vijayan*

Molecular Biophysics Unit, Indian Institute of Science, Bangalore, Karnataka 560012, India

Summary

The crystal structures of jacalin complexed with Gal α-(1,4) Gal and Gal α-(1,3) Gal β-(1,4) Gal have been determined with the primary objective of exploring the effect of linkage on the location of reducing and non-reducing sugars in the extended binding site of the lectin, an issue which has not been studied thoroughly. Contrary to the earlier surmise based on simple steric considerations, the two structures demonstrate that α-linked sugars can bind to jacalin with nonreducing sugar at the primary binding site. This is made possible substantially on account of the hitherto underestimated plasticity of a non-polar region of the extended binding site. Modeling studies involving conformational search and energy minimization, along with available crystallographic and thermodynamic data, indicate a strong preference for complexation with Gal β-(1,3) Gal with the reducing Gal at the primary site, followed by that with Gal α-(1,3) Gal, with the reducing or non-reducing Gal located at the primary binding site. This observation is in consonance with the facility of jacalin to bind mucin type O-glycans containing T-antigen core. © 2016 IUBMB Life, 00(0):000–000, 2016

Keywords: *galactose specific lectin; β-prism I fold; post translational proteolysis; T-antigen binding protein; reducing and non-reducing sugars; mucin type O-glycans*

Introduction

β-prism I as a lectin fold was established in this laboratory through the structure analysis of jacalin, one of the two lectins from jackfruit seeds (1). Jacalin is specific to galactose at the monosaccharide level and to Thomsen–Friedenreich antigen (Gal β-(1,3) GalNAc) at the disaccharide level (2,3). This disaccharide is of non-oncofetal origin and has been shown to have a well documented link to malignancy (4) as it is expressed in more than 85% of human carcinomas such as colon, breast, bladder, buccal cavity, and prostate, as well as on poorly differentiated cells (5–8). Therefore, proteins, which specifically bind to T-antigen, have potential diagnostic value. The crystal structure of the jacalin-methyl α-galactopyranoside complex suggested that a free amino terminus generated by post translational proteolysis, in addition to a stacking interaction involving aromatic residue is important in generating the specificity of jacalin for galactose (1). Artocarpin, the second lectin from jackfruit seeds, is also tetrameric with a β-prism I fold (9,10). Each subunit of artocarpin is made up of a single polypeptide chain without any proteolytic cleavage and is mannose specific. Structures of several galactose specific and mannose specific β-prism I fold lectins and their complexes are now available (11–13). They have nearly the same tertiary structures, but exhibit different modes of quaternary association (14). Protein-sugar interactions at the primary binding site are nearly the same in all the galactose specific β-prism I fold lectins (1,15–22). The same is true about the interactions at the primary site of the mannose specific β-prism I fold lectins (9,10,23–25). However, in both the cases, specificities for oligosaccharides differ among the concerned lectins on account of differences in secondary interactions.

Crystal structures of jacalin with several of these bound sugars have been determined and extensively studied (1,16–19,22). The structure of the tetrameric molecule with

Abbreviations: *Gal, galactose; T-antigen, Thomsen–Friedenreich antigen (Gal β-(1,3) GalNAc); HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; NaCl, sodium chloride; PDB, Protein data bank*

Volume 00, Number 00, Month 2016, Pages 00–00

**Address correspondence to: Mamannamana Vijayan, Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, Karnataka, India. E-mail:mv@mbu.iisc.ernet.in*

Received 5 August 2016; Accepted 23 September 2016

DOI 10.1002/iub.1572

Published online 00 Month 2016 in Wiley Online Library (wileyonlinelibrary.com)

222 symmetry is the same in all the relevant crystals. Each subunit consists of a long α-chain and a short β-chain produced by post translational proteolysis of a single chain precursor. The two chains, however, form an integral part of the prismatic subunit, with three Greek keys arranged around an approximate threefold axis. One of the Greek keys has a break on account of the proteolysis. In spite of its approximate threefold symmetry, each subunit carries only one sugar binding site, which has been characterized thoroughly by the X-ray analysis of several jacalin-carbohydrate complexes (Fig. 1a). Structures of these complexes led to an understanding of the finer details of the interactions of jacalin with different sugars including the T-antigen. However, the nuances of jacalin-carbohydrate interactions remained underexplored in relation to two specific issues. On the basis of the crystal structures of the complexes of jacalin with α-galactosides, it was suggested that the lower affinity of β-galactosides for jacalin is caused by the steric clashes of the β-substituents with the protein (18,19). The difference between the binding of α- and β- galactosides with jacalin was crystallographically explored only recently (22). It turns out that the interactions of an α-galactoside and the corresponding β-galactoside with the lectin remain the same. This is achieved through a distortion of the sugar ring in β-galactosides. Thus, it is not the interaction, but distortion of the molecular geometry that leads to the lower affinity of β-galactosides. This result established distortion of the ligand as a strategy for modulating specificity.

The second underexplored issue was concerned with the influence of glycosidic linkage on the nature of the carbohydrate binding to jacalin. This issue was satisfactorily addressed in the case of mannose/glucose-specific β-prism I fold lectins primarily through crystallographic and modeling studies on appropriate sugar complexes of banana lectin (25). In banana lectin, disaccharides with α-(1,3) linkage prefer to have the nonreducing end at the primary binding site, whereas the reducing end is preferred at the site where the linkage is β-(1,3). Most of the jacalin-disaccharide complexes X-ray analyzed so far have β-(1,3) linkage in the sugar and the reducing end occupies the primary binding site in them (17,18). Melibiose (Gal α-(1,6) Glc) is the only disaccharide for which an α-linked disaccharide complex of jacalin is available (18). The nonreducing end occupies the primary binding site in this complex. The α-(1,6) linkage is highly flexible and it was surmised that it is this flexibility that allows mellibiose to bind to jacalin with the nonreducing end at the primary site. It was also surmised that jacalin cannot bind to α-(1,3) linked disaccharides with the nonreducing end at the binding site on account of steric clashes of the second residue with part of the extended binding site (18). This surmise is yet to be experimentally explored. This work, aimed at exploring this issue and at gaining further information on jacalin-sugar interactions, involves the crystallographic analysis of the complexes of jacalin with an α-(1,4) linked disaccharide and a trisaccharide with α-(1,3) linkage. Furthermore, complementary conformational computations were also carried out. These computations, along with available crystallographic and thermodynamic data, provide additional insights on the influence of glycosidic linkage on the nature of the binding of disaccharides to jacalin.

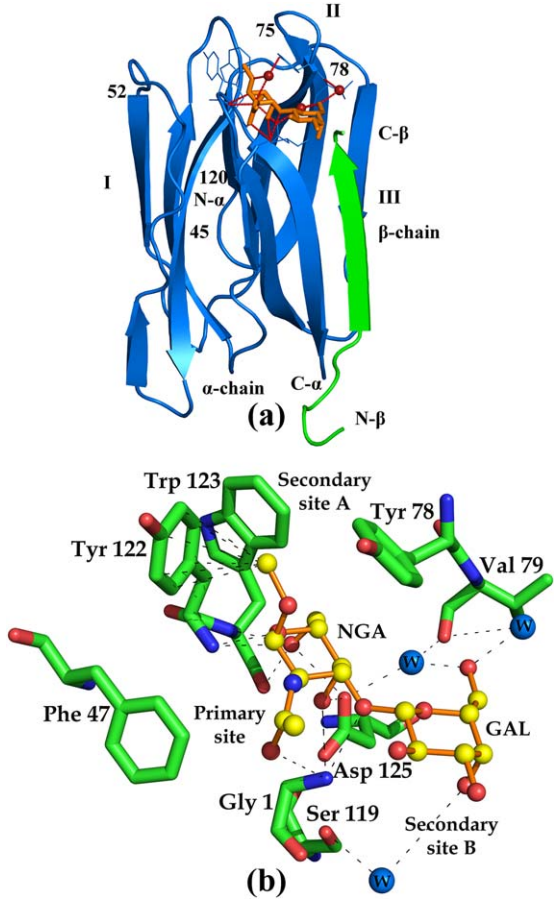

FIG 1

(a) Subunit structure of jacalin bound to Gal β-(1,3) GalNAc-α-OMe. α- and β-chains are indicated in different colors. Individual Greek keys are labeled as I, II, and III. The carbohydrate binding site with the bound sugar is also shown. (b) A detailed view of the sugar-binding site of jacalin in complex with Gal β-(1,3) GalNAc-α-OMe. Invariant water molecules involved in water bridges are shown as spheres. Location of the primary and secondary binding sites A and B are indicated.

Material and Methods

Purification of Jacalin

Jacalin was extracted from the jackfruit seeds by passing the crude seed extract through a galactose cross linked guar gum column by affinity purification and eluting it with galactose in phosphate buffer saline (26). The purified protein was further dialyzed against 20 mM phosphate buffer (pH = 7.3) containing 150 mM NaCl and 0.025% sodium azide to remove all the galactose bound to the protein. Hemagglutination

TABLE 1 ***Data collection and refinement statistics for the jacalin-ligand complexes***

Complex: PBD Code	*Gal α-(1,4) Gal: 5J51*	*Gal α-(1,3) Gal β-(1,4) Gal: 5JM1*
Unit cell dimensions		
a (Å)	58.6	58.5
b (Å)	81.1	81.6
c (Å)	63.3	63.0
$\alpha = \mathfrak{x}$ (degrees)	90	90
β (degrees)	108.1	107.7
Resolution (Highest shell)	35.71–1.67 (1.76–1.67)	46.01–1.95 (2.06–1.95)
No. of observations	225,458	139,602
No. of unique reflections	63,612	38,922
Completeness (%)	96.9 (93.7)	94.8 (91.0)
$\langle I/\sigma\ (I)\rangle$	10.9 (2.0)	8.2 (3.7)
CC1/2	0.667	0.366
R_{merge} (%)[a]	8.3 (63.9)	13.2 (33.7)
Multiplicity	3.5 (3.4)	3.6 (3.3)
***R*-factor (%)**	16.4 (27.4)	18.2 (25.9)
R_{free} (%)[b]	21.3 (31.9)	23.7 (31.3)
No. of atoms		
Proteins	4564	4544
Ligands	189	94
Water O atoms	466	403
Rms dev. from ideal values		
Bond lengths (Å)	0.020	0.018
Bond angles (degrees)	1.992	1.897
Chiral volume	0.153	0.115
Average *B* factors (Å^2)		
Overall	16.9	13.2
Protein	14.9	12.1
Ligands	32.7	32.6
Water O atoms	30.3	20.6
Ramachandran plot		
Core region (%)	89.6	90.0
Additionally allowed region (%)	10.2	9.5
Generously allowed region (%)	0.2	0.4
Disallowed region (%)	0	0

The crystals belong to space group $P2_1$ and contain a crystallographically independent tetramer.

[a] $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, *where* $I_i(hkl)$ *is the* i*th intensity measurement of a reflection,* $\langle I(hkl)\rangle$ *is the average intensity value of that reflection and the summation is over all measurements.*

[b] *10% of the reflections were used for the* R_{free} *calculations.*

measurements were also carried out after every batch of purification to ascertain the activity of the purified batch. Protein concentrations were checked using spectrophotometry.

Crystallization

Native crystals of jacalin were grown in a condition similar to that used in the crystallization mentioned earlier (16). Vapor diffusion technique was employed at 25 °C by equilibrating a 4 μL drop of 12 mg/mL protein in 10 mM HEPES buffer (pH 7.4) containing 150 mM NaCl, mixed with 4 μL of the reservoir solution containing 15% poly(ethylene glycol) 8000, 10% (v/v) isopropanol and 100 mM HEPES (pH 7.4). Crystals of approximate dimensions $0.1 \times 0.05 \times 0.05$ mm^3 grew in about 4 weeks. Since co-crystallization attempts were not successful, soaking experiments were conducted to obtain the jacalin-ligand crystal complexes. The Gal α-(1,4) Gal and Gal α-(1,3) Gal β-(1,4) Gal sugars were purchased from Sigma Aldrich and Dextra, respectively. Typically, 30 times molar excess of the ligand in the mother liquor was used for soaking. The soaking time was 48 h in both the cases.

Data Collection and Processing

Data from the complexes were collected at home source using a MAR345 image plate detector mounted on a Bruker MicroStar rotating anode X-ray generator at 100 K using 25% ethylene glycol as the cryoprotectant. The intensity data were processed and merged using iMosflm (27) and scaled with SCALA (28) in the CCP4 program suite (29). The intensity data were converted into structure-factor amplitudes using TRUNCATE (30) in the CCP4 suite. The data collection statistics along with the cell parameters are given in Table 1.

Structure Refinement and Validation

The structures were refined using REFMAC (31) in CCP4 and model building was carried out using Coot v0.7.1 (32) with the coordinates provided in PDB for native jacalin (PDB Code: 1KU8) (16) as the starting model. Addition of sugar ligands and water O atoms was commenced using PRODRG (33) when the R and $R_{\rm free}$ were close to 18 and 23% respectively. The water O atoms were located based on peaks with heights $>1.0\sigma$ in $2F_o - F_c$ and 3σ in $F_o - F_c$ maps. The possibility of alternate ligand conformations were also evaluated before finalizing on the ligand fitting. The refined models were validated using PROCHECK (34) and the MOLPROBITY (35) web server. Refinement statistics are also summarized in Table 1. The refinement of the ligand structure was validated by calculating the electron density using simulated annealing omit maps generated by Crystallography and NMR System (CNS; (36)) contoured at 3σ.

PDB References

Jacalin in complex with Gal α-(1,4) Gal, 5J51 and in complex with Gal α-(1,3) Gal β-(1,4) Gal, 5JM1.

Analysis of Structures

Structure alignments were carried out using ALIGN (37). All pictorial illustrations were generated using PyMOL (38).

Software used for analyzing structures includes Swiss PDB viewer (39), Mercury (40), UCSF Chimera (41), and Privateer (42).

Molecular Modeling and Energy Calculations

A locally written script embedded in UCSF CHIMERA (41) was used for generating sterically acceptable different conformers of the free disaccharides. Models of the complexes were generated substantially on the basis of crystallographic information, as described in Results and Discussion. An automated version of ACPYPE (AnteChamber PYthon Parser interfacE) (43) running ANTECHAMBER (44) was employed to generate carbohydrate topologies compatible with GROMACS v3.3.1 (45). Inconsistencies of these topology files when compared to the standard files obtained from GLYCAM06 (46) were examined and corrected. The inconsistencies pertained to torsion angles, partial atomic charges, and atom name conventions. The simulation box was generated using the editconf module of GROMACS with the criterion that the minimum distance between the solute and edge of the box was at least 7.5 Å. The protein models were solvated with TIP4P water model using program genbox available in the GROMACS suite. Sodium or chloride ions were added to neutralize the overall charge of the system. The energy minimization of the complexes was then performed using GROMACS v3.3.1 (45) running on parallel processors with the OPLS-AA/L force field (47) using 200 cycles of steepest descent. The potential energy values converged within 200 cycles in every case. The lectin-sugar binding energy in these models was calculated subsequently by using the python script write_component_energies.py implemented under MGL Tools 4.6 (48) implementing Autodock 4.1 (49). To ensure that there is no change in pose of the sugar bound to the jacalin, the energy values obtained from this script was used in the analysis without implementing a conventional docking routine.

Results and Discussion

Overall Features

Each site of jacalin is made up of loops 46–52 (from Greek key I), 76–82 (Greek key II), and 122–125 (Greek key III) and the N terminus of the α-chain. The extended binding site of jacalin comprising of the secondary binding site A, the primary site and secondary binding site B forms a contiguous patch for the sugars to bind. The extended binding site and protein-carbohydrate interactions in a typical jacalin-sugar complex (18) are illustrated in Fig. 1b. The primary binding site is made up of the side chains of Phe 47, Tyr 78, and Asp 125, the backbone nitrogen and oxygen atoms of Tyr 122 and Trp 123 and the free amino group of Gly 1 of the α-chain. Among these, Tyr 78 stacks against the galactose ring of the ligand. The aromatic side chains of Tyr 78, Tyr 122, and Trp 123 constitute the secondary binding site A, which is wholly hydrophobic. The ligand is involved only in water mediated interactions with the secondary site B composed of the main chain nitrogen and oxygen atoms of Val 79, Ser 119 OG, and the carboxyl terminal region of the β-chain.

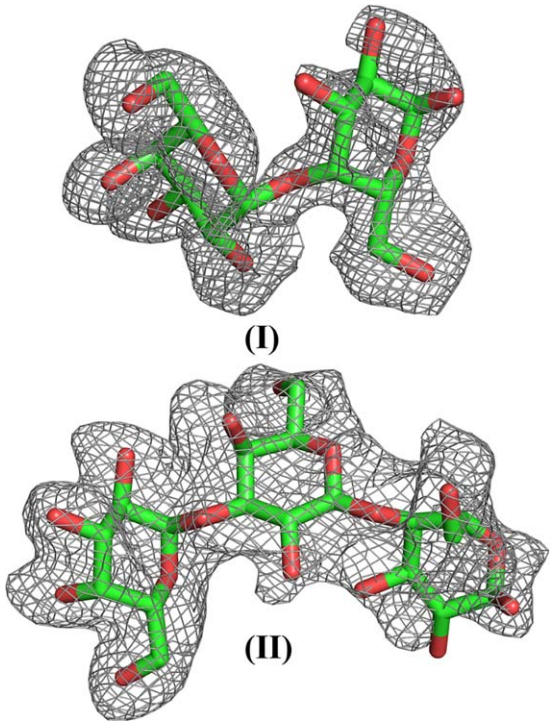

FIG 2 *Electron densities in simulated annealing omit maps contoured at 3σ for (I) Gal α-(1,4) Gal and (II) Gal α-(1,3) Gal β-(1,4) Gal in the respective complexes.*

The sugar complexes of jacalin reported here were prepared by soaking native crystals of jacalin in solutions containing Gal α-(1,4) Gal (I) and Gal α-(1,3) Gal β-(1,4) Gal (II). The crystal structures of the complexes were determined at medium to high resolutions (Table 1). The accessibility of the four binding sites in the crystal structure exhibits differences. Furthermore, possibilities of additional protein-sugar interactions involving symmetry related molecules exist at some sites (22). Consequently, all the sites are not occupied by the same ligand in all the cases. Subunit C is fully occupied by both the sugars. Therefore, subunit C forms the basis for the discussing the sugar binding. The difference densities in the simulated annealing omit maps contoured at 3σ for bound sugars are shown in Fig. 2.

Location and Interactions of the Sugar Molecules

An interesting aspect of lectin-sugar interactions in the α- and β-linked disaccharide bound jacalin complexes is the location of the reducing and nonreducing sugars in them at different regions of the extended binding site. In all the complexes involving β-(1,3) linked disaccharides, the primary binding site is always occupied by the reducing sugar (17,18) with the nonreducing sugar occupying the secondary binding site B (Fig. 1b). Also, the substituents at the anomeric oxygen of the reducing sugar, a methyl group in most cases, point toward and interacts with the hydrophobic secondary site A.

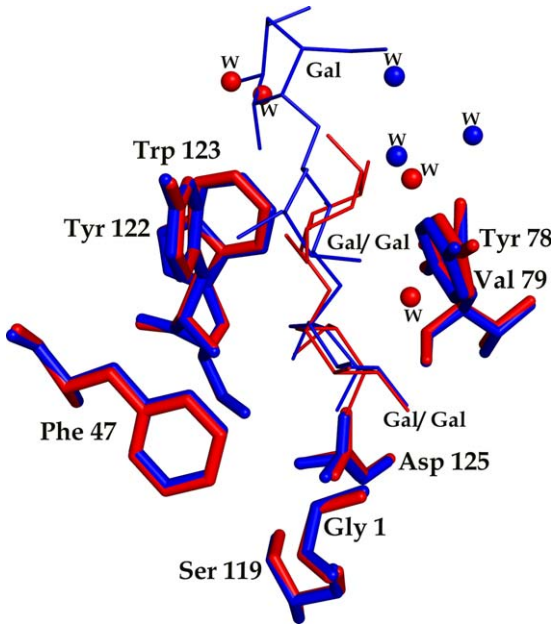

FIG 3 *Location of the bound Gal α-(1,4) Gal (red) and Gal α-(1,3) Gal β-(1,4) Gal (blue) in the respective complexes. The bound waters are shown as spheres.*

Interactions of the methyl group with the subsite have been discussed earlier (1,18). In I, the disaccharide has an α-(1,4) linkage. In the only trisaccharide (II) which has been studied in complex with jacalin, the first and the second sugars are connected through an α-(1,3) linkage. The nonreducing sugar is located at the primary binding site in both the complexes (Fig. 3) while secondary binding site B remains unoccupied. In addition, the second sugar residue is located at secondary binding site A in both the complexes. In the trisaccharide complex, the third sugar residue points into the solution. This residue has a half boat conformation (42,50) unlike all other sugar residues in the structures, which have the expected chair conformation (42). The location of α-linked sugars in the two complexes and the side chains of aromatic amino acids in secondary site A are similar to those in mellibiose (Gal α-(1,6) Glc). Furthermore, the reducing galactose in Gal α-(1,3) Gal β-(1,4) Gal is anchored on secondary site A by forming a stacking interaction with Tyr 122. A stacking interaction is defined by the distance between the center of the two rings (N in Å) and the angle between the two planar groups (θ in degrees). The values of N and θ in this case are 4.1 Å and 16°, respectively.

Secondary site A presents an interesting picture in terms of its plasticity and in relation to its effect on the geometry of the ligand. As noted earlier, the side chains of Tyr 78 and Tyr 122, particularly that of Tyr 78, assumes different conformations in the four subunits of the native structure, indicating their intrinsic flexibility (16,22). The flexibility is substantially reduced when ligands with substituents pointing to these sites bind to jacalin. However, the residual plasticity of the side chains is used to accommodate different kinds of substituents. The extent of the flexibility is also evident in the movement of Tyr 78 OH group in addition to the variability in the side chain conformation angles of the aromatic residue. The oxygen atom moves by 1.3 Å in the mellibiose complex with respect to its structure in the galactose complex (18). The corresponding movements in the Gal α-(1,3) Gal β-(1,4) Gal and Gal α-(1,4) Gal complexes are 1.7 and 2.2 Å, respectively. The steric resistance offered by the amino acids at secondary site A also leads to the distortion of the ligand, where needed. Thus, it would appear that the plasticity of secondary site A and the distortion of the ligand molecule are responsible for relieving the unfavorable steric clashes.

Complexes of jacalin with β-linked disaccharides involved the identification of a route to the binding site through the secondary site B (Fig. 4). The complexes with α-linked sugars reported here, particularly that with the trisaccharide, delineated a route through the secondary site A. Thus, the two sets of complexes, between them, indicate how oligosaccharides with β- and α- linkages in the terminal residues can enter the binding site of jacalin (Fig. 4).

Modeling and Energetic Studies

Structural studies described above showed the earlier surmise that jacalin cannot bind to α-(1,3) linked disaccharides with the nonreducing end at the primary binding site, to be incorrect. Indeed, the effect of the glycosidic linkage on the nature of sugar binding to jacalin has to be systematically explored. Therefore, such an exploration using modeling on the basis of crystallographic results was undertaken. The protein component of a subunit of the jacalin-methyl T-antigen complex (18) in which the primary binding site as well as the two secondary

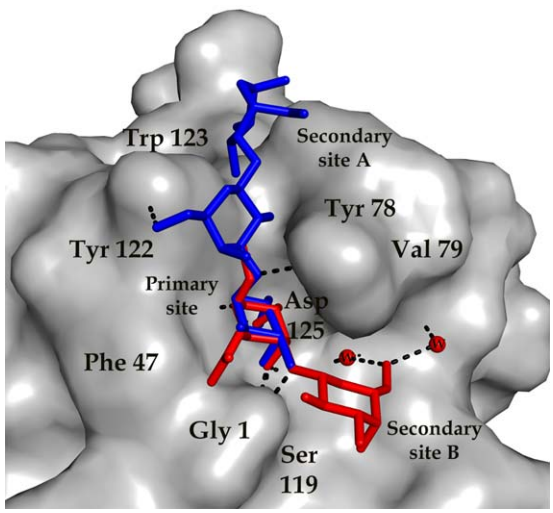

FIG 4 *Locations of a typical β-linked sugar Gal β-(1,3) GalNAc-α-OMe (red) and α-linked Gal α-(1,3) Gal β-(1,4) Gal (blue) in the extended binding site of jacalin shown in van der Waals representation (gray). Only conserved intrasubunit waters contributing to the ligand binding have been shown as spheres for clarity.*

TABLE 2 *Location of residues in the galactobioside ligands in the modeled complexes*

S. No.	Disaccharide type	Primary binding site	Secondary site A	Secondary site B	No. of allowed conformers	Energy range (kJ/mol)
(a)	Gal α-(1,3) Gal	Nonreducing Gal	Reducing Gal	–	21	−27.0 to −22.7
(b)	Gal α-(1,3) Gal	Reducing Gal	–	Nonreducing Gal	79	−27.8 to −18.3
(c)	Gal β-(1,3) Gal	Reducing Gal	–	Non-reducing Gal	108	−26.6 to −17.8
(d)	Gal β-(1,3) Gal	Nonreducing Gal	Reducing Gal	–	12	−26.1 to −22.6
(e)	Gal α-(1,4) Gal	Nonreducing Gal	Reducing Gal	–	10	−27.0 to −22.8
(f)	Gal α-(1,4) Gal	Reducing Gal	–	Nonreducing Gal	0	–
(g)	Gal β-(1,4) Gal	Reducing Gal	–	Nonreducing Gal	0	–
(h)	Gal β-(1,4) Gal	Nonreducing Gal	Reducing Gal	–	5	−26.8 to −24.8

The number of allowed conformers in each case and the range of interaction energies in them are also given.

binding sites are occupied, was used in the modeling. Initial models of Gal α-(1,3) Gal, Gal β-(1,3) Gal, Gal α-(1,4) Gal, and Gal β-(1,4) Gal were constructed using standard bond lengths and angles. Each disaccharide model was then positioned on to the jacalin molecule using the known jacalin-galactose interactions in two ways, one with the reducing Gal at the primary binding site and the other with the nonreducing Gal at the primary binding site. The resulting locations of the two sugars in the eight models are indicated in Table 2.

The natural orientation of the two sugar residues in the disaccharide is defined by the two glycosidic linkage torsion angles φ [O5-C1-O1-C3′, where the linkage is (1,3) or O5-C1-O1-C4′, where the linkage is (1,4)] and ψ (C1-O1-C3′-C4′ or C1-O1-C4′-C5′). The orientation of the alcoholic group in each residue is defined by the torsional angle ω (O6-C6-C5-O5 and O6′-C6′-C5′-O5′). Of the two ω angles, the one in the residue at primary binding site is fixed through hydrogen bonded interactions and hence is not variable. Thus, the conformations of the disaccharide in the complex is determined by the three torsion angles φ, ψ, and ω of the sugar at the secondary site. To start with, φ and ψ were rotated at 30° intervals leading to 144 conformations in each case. ω has preferred values of −60°, 180°, and +60°. For each combination of φ and ψ angles, three conformers corresponding to the three preferred values of ω were generated. Thus, 432 independent conformations were generated for the free saccharide in each case. Of these, those containing intermolecular or intramolecular contact distances <2.2 Å were rejected as sterically unacceptable. Complexes involving the remaining models were energy minimized using the procedures described in Material and Methods. The energy minimized models were again carefully examined for short contacts and those containing contact distances <2.5 Å were rejected. The remaining ones were accepted as possible models. The numbers of such models for each of the eight complexes are given in Table 2 along with the range of interaction energies in them. The distribution of φ, ψ angles in the acceptable models is illustrated in Fig. 5. As expected, the values of φ and ψ of the β-linked as well as the α-linked disaccharides observed in the crystal structures are among the possible conformations indicated by the calculations.

Results of the modeling studies need to be treated with caution. This is particularly true about the energies that result from the calculations. However, the results obtained above provide important insights into the effect of the nature of the linkage on the orientation of the sugar ligand with respect to the primary binding site of the lectin. They indicate that the binding site is capable of accommodating, with varying degrees of facility, different linkages with reducing or nonreducing sugar at the primary binding site, although the nature of interactions at the primary and secondary binding sites remains substantially unaltered. The totally disallowed arrangements are that of Gal α-(1,4) Gal and Gal β-(1,4) Gal with reducing Gal at the primary binding site. In both the cases, the nonreducing Gal severely clashes with the residues at the primary site and secondary site B. For both the sugars, an arrangement with the nonreducing Gal at the primary site is acceptable, though not favored as evidenced by the smaller number of sterically allowed conformations that these sugars can assume when binding to jacalin. (1,3)-linked galactobiosides can bind to jacalin irrespective of the anomeric nature of the linkage. They can bind with the reducing or the nonreducing sugar at the primary binding site. However, numbers and energy ranges indicate that the arrangement with the reducing Gal at the primary site is preferred only when the linkage is β. It would thus appear that jacalin is meant predominantly for (1,3)-linked galactobioses with the reducing sugar at the primary site. Even among them, that with the β-linkage is preferred.

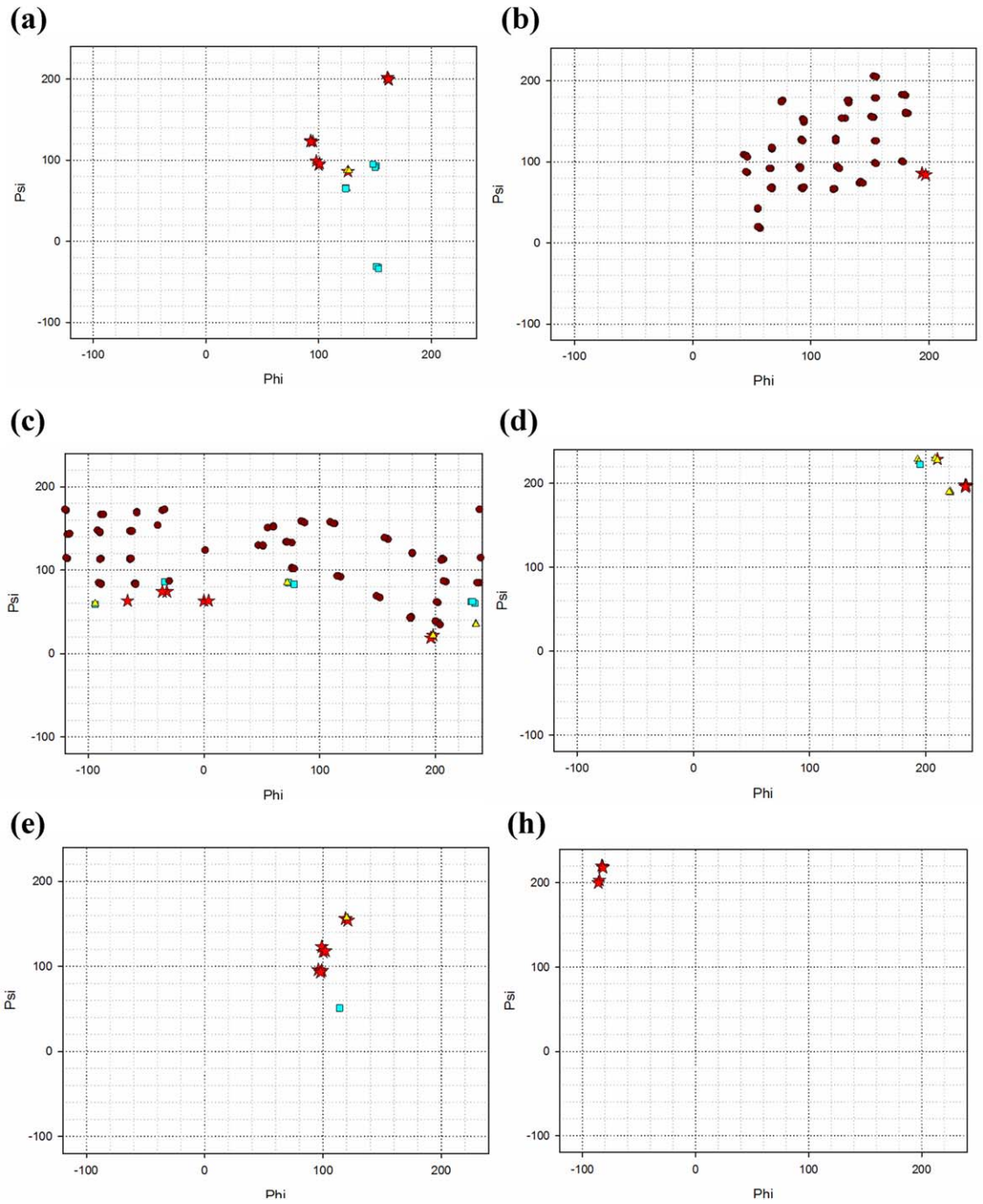

FIG 5

Distribution of φ, ψ angles in the acceptable models belonging to the eight complexes involving both (1,3)- and (1,4)-linked galactosbiosides as given in Table 2. The same serial number of the combination as given in Table 2 has been used in the figure for clarity. The conformations are energetically ranked at different contour levels, $\Delta(\Delta G_{bind}) < 2$ as stars, $\Delta(\Delta G_{bind}) < 3$ as triangles, $\Delta(\Delta G_{bind}) < 4$ as squares and $\Delta(\Delta G_{bind})$ between 4 and 8 as circles.

The above observations are in consonance with the available results of thermodynamic and crystallographic studies. Disaccharides involving Gal and GalNAc exhibit maximum affinity to jacalin where the linkage is β-(1,3) and when there is no β-substitution at the second sugar residue (18). In all the crystal structures of jacalin involving such β-(1,3) linked disaccharides, the reducing sugar occupies the primary binding site (17,18). Modeling indicates that the largest number of conformations for the disaccharide is allowed where the linkage is β-(1,3) and the reducing sugar is at the primary site (Table 2). A location with the nonreducing sugar at the site is also allowed for the β-(1,3) linked disaccharides but the one

with the reducing sugar is decisively favored. Modeling also suggests that α-(1,3) galactobiose can also bind jacalin, though not as well as the β-(1,3) counterpart can. This is in agreement with the thermodynamic measurements done earlier (51). However, the choice of reducing or nonreducing sugar occupying the primary binding site is somewhat ambiguous in terms of the energies and possible number of conformations in the α-(1,3) linked galactobiose. Thermodynamic measurements indicates a good binding for the Gal α-(1,3) Gal disaccharide (51) but cannot obviously distinguish between the two modes of binding in terms of the location of the reducing or nonreducing sugar at the binding site. One of the jacalin complexes reported here is an α-(1,3) linked trisaccharide (Gal α-(1,3) Gal β-(1,4) Gal) bound to the lectin. The trisaccharide binds to jacalin with non-reducing sugar at the primary binding site. The choice is presumably influenced by the presence of the additional β-substituted sugar. Gal α-(1,4) Gal binds to jacalin, but with much less affinity than disaccharides with β-(1,3) and α-(1,3) disaccharides (51). Modeling suggests that Gal α-(1,4) Gal can bind to jacalin only with the nonreducing sugar in the primary site. That is what is indeed observed in the crystal structure of the jacalin complex of Gal α-(1,4) Gal reported here. Gal β-(1,4) Gal does not show detectable binding to jacalin (unpublished results), a result again in consonance with that obtained from modeling.

We had demonstrated earlier that four of the five mucin type O-glycans (which contain the T-antigen core) with this preferred β-(1,3) linkage with the reducing sugar at the primary binding site can bind to jacalin (18). However, attempts to dock the disialyated hexasaccharide NeuNAc α-(2,3) Gal β-(1,3) (NeuNAc α-(2,3) Gal β-(1,4) GlcNAc β-(1,6)) GalNAc, in a similar manner to jacalin resulted in severe steric clashes. The hexasaccharide can be now docked without serious clashes with the nonreducing Gal at the primary binding site. However, the binding of Gal β-(1,3) Gal with nonreducing Gal at the primary binding site is energetically less favorable than that of the same disaccharide with the reducing Gal at the primary binding site (Table 2). That explains the apparent inability of the O-mucin with the hexasaccharide core to bind to jacalin (52,53).

Conclusions

Crystal structures of the two α-linked sugars provide new insights into the interactions of carbohydrates with the lectin. In the complex with Gal α-(1,4) Gal, the nonreducing sugar occupies the primary binding site and the reducing sugar occupies secondary site A. In the complex involving the trisaccharide Gal α-(1,3) Gal β-(1,4) Gal, the nonreducing sugar is located at the primary binding site and the middle residue in secondary site A. This is not in consonance with the earlier surmise that the nonreducing sugar cannot occupy the primary site in α-(1,3) linked galactobiosides that interact with jacalin. Analysis of the two structures reported here indicate that the reason for accommodating the α-(1,3) and α-(1,4) linked galactobiosides with nonreducing sugar located at the jacalin primary binding site is on account of the inherent flexibility of the carbohydrate binding site with special reference to Tyr78.

The available crystallographic and thermodynamic data on jacalin-sugar interactions are in consonance with the results of modeling studies involving conformational search and energy minimization, carried out with the complexes involving Gal α-(1,3) Gal, Gal β-(1,3) Gal, Gal α-(1,4) Gal, and Gal β-(1,4) Gal. The combined results indicate a strong preference for Gal β-(1,3) Gal for complex formation with the reducing Gal at the primary binding site followed by that for Gal α-(1,3) Gal, with the reducing or the nonreducing Gal at the primary binding site. Gal β-(1,3) Gal with the nonreducing sugar at the primary site can probably bind weakly. (1–4) Linked disaccharides can also probably bind to jacalin very weakly with nonreducing sugar at the primary binding site. Such disaccharides cannot bind to the lectin at all with the reducing Gal at the primary site. Indeed, the strong preference of jacalin for Gal β-(1,3) Gal with the reducing Gal in the primary binding site explains the facility of jacalin to bind to mucin type O-glycans containing the T-antigen core.

Acknowledgements

The datasets were collected at the X-ray facility for Protein Crystal Structure Determination and Protein Design at this Institute, supported by the Department of Science and Technology (DST). Part of the computations was carried out at the Graphics Facility supported by DBT. K. V. A. is a Council of Scientific and Industrial Research (CSIR) Senior Research Fellow. A. S. is a CSIR Bhatnagar Fellow and M. V. is Albert Einstein Professor of the Indian National Science Academy (INSA). The work is supported by a grant from the Science and Engineering Research Board (SERB).

References

[1] Sankaranarayanan, R., Sekar, K., Banerjee, R., Sharma, V., Surolia, A., et al. (1996) A novel mode of carbohydrate recognition in jacalin, a *Moraceae* plant lectin with a β-prism fold. Nat. Struct. Biol. 3, 596–603.

[2] Sastry, M. V., and Surolia, A. (1986) Intrinsic fluorescence studies on saccharide binding to *Artocarpus integrifolia* lectin. Biosci. Rep. 6, 853–860.

[3] Sastry, M. V., Banarjee, P., Patanjali, S. R., Swamy, M. J., Swarnalatha, G. V., and Surolia, A. (1986) Analysis of saccharide binding to *Artocarpus integrifolia* lectin reveals specific recognition of T-antigen (β-D-Gal-(β-(1-3) D-GalNAc). J. Biol. Chem. 261, 11726–11733.

[4] Springer, G. F. (1984) T and Tn, general carcinoma auto antigens. Science 224, 1198–1206.

[5] Thatcher, N., Hashmi, K., Chang, J., Swindell, R., and Crowther, D. (1980) Anti-T antibody in malignant melanoma patients. Influence of response and survival following chemotherapy-changes in serum levels following *C. parvum*, BCG immunization. Cancer 46, 1378–1382.

[6] Vos, G. H., and Brain, P. (1981) Heterophile antibodies, immunoglobulin levels, and the evaluation on anti-T activity in cancer patients and controls. S. Afr. Med. J. 60, 133–136.

[7] Bray, J., Maclean, G. D., Dusel, F. J., and McPherson, T. A. (1982) Decreased levels of circulating lytic anti-T in the serum of patients with metastatic gastro-intestinal cancer: a correlation with disease burdan. Clin. Expt. Immunol. 47, 176–182.

[8] Springer, G. F., Desai, P. R., Wise, W., Carlstedt, S. C., Tegtmeyer, H., et al. (1990) Pancarcinoma T and Tn epitopes: autoimmunogens and diagnostic markers that reveal incipient carcinomas and help establish prognosis. Immunol. Ser. 53, 587–612.
[9] Pratap, J. V., Jeyaprakash, A. A., Rani, P. G., Sekar, K., Surolia, A., et al. (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.
[10] Jeyaprakash, A. A., Srivastav, A., Surolia, A., and Vijayan, M. (2004) Structural basis for the carbohydrate specificities of artocarpin: variation in the length of a loop as a strategy for generating ligand specificity. J. Mol. Biol. 338, 757–770.
[11] Chandran, T., Sharma, A., and Vijayan, M. (2013) Generation of ligand specificity and modes of oligomerization in β-prism I fold lectins. Adv. Protein Chem. Struct. Biol. 92, 135–178.
[12] Abhinav, K. V., and Vijayan, M. (2014) Structural diversity and ligand specificity of lectins. The Bangalore effort. Pure Appl. Chem. 86, 1335–1355.
[13] Pérez, S., Sarkar, A., Rivet, A., Breton, C., and Imberty, A. (2015) Glyco3D: a portal for structural glycosciences. Methods Mol. Biol 1273, 241–258.
[14] Sharma, A., and Vijayan, M. (2011) Quaternary association in β-prism I fold plant lectins: insights from X-ray crystallography, modelling and molecular dynamics. J. Biosci. 36, 793–808.
[15] Lee, X., Thompson, A., Zhang, Z., Ton-that, H., Biesterfeldt, J., et al. (1998) Structure of the complex of *Maclura pomifera* agglutinin and the T-antigen disaccharide, Gal β-(1,3) GalNAc. J. Biochem. 273, 6312–6318.
[16] Bourne, Y., Astoul, C. H., Zamboni, V., Peumans, W. J., Menu-Bouaouiche, L., et al. (2002) Structural basis for the unusual carbohydrate-binding specificity of jacalin towards galactose and mannose. Biochem. J. 364, 173–180.
[17] Jeyaprakash, A. A., Rani, P. G., Reddy, G. B., Banumathi, S., Betzel, C., et al. (2002) Crystal structure of the jacalin-T-antigen complex and a comparative study of lectin-T-antigen complexes. J. Mol. Biol. 321, 637–645.
[18] Jeyaprakash, A. A., Katiyar, S., Swaminathan, C. P., Sekar, K., Surolia, A., and Vijayan, M. (2003) Structural basis of the carbohydrate specificities of jacalin: an X-ray and modeling study. J. Mol. Biol. 332, 217–228.
[19] Jeyaprakash, A. A., Jayashree, G., Mahanta, S. K., Swaminathan, C. P., Sekar, K., et al. (2005) Structural basis for the energetics of jacalin-sugar interactions: promiscuity versus specificity. J. Mol. Biol. 347, 181–188.
[20] Sharma, A., Sekar, K., and Vijayan, M. (2009) Structure, dynamics, and interactions of jacalin. Insights from molecular dynamics simulations examined in conjunction with results of X-ray studies. Proteins 760–777.
[21] Huang, J., Xu, Z., Wang, D., Ogata, C. M., Palczewski, K., et al. (2010) Characterization of the secondary binding sites of *Maclura pomifera* agglutinin by glycan array and crystallographic analyses. Glycobiology 20, 1643–1653.
[22] Abhinav, K. V., Sharma, K., Swaminathan, C. P., Surolia, A., and Vijayan, M. (2015) Jacalin-carbohydrate interactions: distortion of the ligand molecule as a determinant of affinity. Acta Crystallogr. D Biol. Crystallogr. 71, 324–331.
[23] Meagher, J. L., Winter, H. C., Ezell, P., Goldstein, I. J., and Stuckey, J. A. (2005) Crystal structure of banana lectin reveals a novel second sugar binding site. Glycobiology 15, 1033–1042.
[24] Singh, D. D., Saikrishnan, K., Kumar, P., Surolia, A., Sekar, K., and Vijayan, M. (2005) Unusual sugar specificity of banana lectin from *Musa paradisiaca* and its probable evolutionary origin. Crystallographic and modelling studies. Glycobiology 60, 1025–1032.
[25] Sharma, A., and Vijayan, M. (2011) Influence of glycosidic linkage on the nature of carbohydrate binding in β-prism I fold lectins: an X-ray and molecular dynamics investigation on banana lectin-carbohydrate complexes. Glycobiology 21, 23–33.
[26] Kumar, G. S., Appukuttan, P. S., and Basu, D. (1982) α-D-galactose-specific lectin from jack fruit (*Artocarpus integra*) seed. J. Biosci. 4, 257–261.
[27] Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. W. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. D Biol. Crystallogr. 67, 271–281.
[28] Evans, P. R. (1993) Data reduction. Proceedings of CCP4 Study Weekend, 1993, on Data Collection and Processing, Daresbury Laboratory, Warrington, UK. pp. 114–122.
[29] Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242.
[30] French, S., and Wilson, K. (1978) On the treatment of negative intensity observations. Acta Crystallogr. A Crystallogr. 34, 517–525.
[31] Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. 67, 355–367.
[32] Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.
[33] Schuttelkopf, A. W., and van Aalten, D. M. F. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr. D Biol. Crystallogr. 60, 1355–1363.
[34] Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26, 283–291.
[35] Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 35, 375–383.
[36] Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR system. Nat. Protocols 2, 2728–2733.
[37] Cohen, G. H. (1997) ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. J. Appl. Cryst. 30, 1160–1161.
[38] The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC.
[39] Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., et al. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 42, 252–258.
[40] Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., et al. (2008) Mercury CSD 2.0 - New features for the visualisation and investigation of crystal structures. J. Appl. Cryst. 41, 466–470.
[41] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., et al. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
[42] Agirre, J., Iglesias-Fernández, J., Rovira, C., Davies, G. J., Wilson, K. S., et al. (2015) Privateer: software for the conformational validation of carbohydrate structures. Nat. Struct. Biol. 22, 833–834.
[43] Sousa da Silva, A. W., and Vranken, W. F. (2012) ACPYPE - AnteChamber PYthon Parser interfacE. BMC Res. Notes 5, 367.
[44] Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. J. Mol. Graph. Model 25, 247–260.
[45] Lindahl, E., Hess, B., and van der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. J. Mol. Model. 7, 306–317.
[46] Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., et al. (2008) GLYCAM06: a generalizable biomolecular force field Carbohydrates. J. Comput. Chem. 29, 622–655.
[47] Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. (2001) Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. J. Phys. Chem. B 105, 6474–6487.
[48] Sanner, M. F. (1999) Python: A Programming Language for Software Integration and Development. J. Mol. Graph. Model 17, 57–61.
[49] Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., et al. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J. Comput. Chem. 2785–2791.
[50] Cremer, D., and Pople, J. A. (1975) General definition of ring puckering coordinates. J. Am. Chem. Soc. 97, 1354–1358.
[51] Mahanta, S. K., Sastry, M. V. K., and Surolia, A. (1990) Topography of the combining region of a Thomsen-Friedenreich-antigen-specific lectin, jacalin [*Artocarpus integrifolia* (jackfruit) agglutinin]. A thermodynamic and circular-dichroism spectroscopic study. Biochem. J. 265, 831–840.
[52] Meaumura, K., and Fukuda, M. (1992) Poly-N-acetyllactosaminyl O-glycans attached to leukosialin. The presence of sialyl Le(x) structures in O-glycans. J. Biochem. 34, 24379–24386.
[53] Tachibana, K., Nakamura, S., Wang, H., Iwasaki, H., Tachibana, K., et al. (2006) Elucidation of binding specificity of Jacalin toward O-glycosylated peptides: quantitative analysis by frontal affinity chromatography. Glycobiology 16, 46–53.