Glycobiology, 2017, vol. 27, no. 12, 1120–1133
doi: 10.1093/glycob/cwx081
Advance Access Publication Date: 3 October 2017
Original Article

Structural Biology

Distinct roles for each *N*-glycan branch interacting with mannose-binding type Jacalin-related lectins Orysata and Calsepa

Masamichi Nagae[2,5,6], Sushil K Mishra[2,6], Shinya Hanashima[3], Hiroaki Tateno[4], and Yoshiki Yamaguchi[2,1]

[2]Structural Glycobiology Team, Systems Glycobiology Research Group, RIKEN Global Research Cluster, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan, [3]Department of Chemistry, Osaka University, Machikaneyama, Toyonaka, Osaka 560-0043, Japan, and [4]National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan

[1]To whom correspondence should be addressed: Tel: +81-48-467-9619; Fax: +81-48-467-9620; e-mail: yyoshiki@riken.jp

[5]Present address: Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[6]These authors equally contributed to this work.

Received 30 April 2017; Revised 2 September 2017; Editorial decision 2 September 2017; Accepted 6 September 2017

Abstract

Mannose-binding type Jacalin-related lectins (mJRLs) bind to branched *N*-glycans via conserved sugar-binding sites. Despite, significant 3D structural similarities, each mJRL is known to have a unique binding preference toward various *N*-glycans. However, the molecular basis of varying binding preference is substantially unknown. Here, we report a detailed comparison of *N*-glycan-binding preference for two mJRLs, Orysata and Calsepa using frontal affinity chromatography (FAC), X-ray and molecular modeling. The FAC analysis using a panel of *N*-glycans shows difference in *N*-glycan-binding preference between the lectins. Orysata shows broader specificity toward most high-mannose-type glycans as well as biantennary complex-type glycans bearing an extension on the Manα1–6 branch. Whereas, Calsepa shows narrow specificity to complex-type glycans with bisecting GlcNAc. The X-ray crystallographic structure reveals that two Orysata lectins bind to one biantennary *N*-glycan (2:1 binding) where one lectin binds to mannose of the α1–3 branch, while the other interacts with an *N*-acetylglucosamine of the α1–6 branch. In contrast, Calsepa shows 1:1 binding where α1–3 branch and core chitobiose region *N*-glycan interacts with lectin, while α1–6 branch is flipped-back to the chitobiose core. Molecular dynamics study of Orysata bound to *N*-glycans substantiate possibility of two-binding modes for each *N*-glycan. Binding free energies calculated separately for α1–3 and α1–6 branches of each *N*-glycan suggest both branches can bind to Orysata. Overall these results suggest that each branch of *N*-glycan has a distinct role in binding to mJRLs and the nonbinding branch can contribute significantly to the binding affinity and hence to the specificity.

Key words: branched glycan, frontal affinity chromatography, mannose-binding type Jacalin-related lectin, molecular dynamics simulation, X-ray crystallography

Introduction

Accumulating evidence shows that glycan–protein interactions play various physiological roles, such as pathogen recognition, cell–cell interactions and quality control of nascent glycoproteins (Drickamer and Taylor 2015; Nagae and Yamaguchi 2015). Detailed investigation of the 3D structure, dynamics and interaction mode of glycans is essential in order to gain a better understanding of such physiological events. A large number of lectin-sugar complex structures have been reported, and most studies have focused on lectins that recognize mono- and disaccharides located on the termini of glycans. For example, two animal lectins, siglecs and asialoglycoprotein receptor, interact with *N*-glycans by recognizing the terminal sialic acid and galactose, respectively (Nagae and Yamaguchi 2014).

In contrast, several lectins specifically bind the internal unit of branched *N*-glycans. C-type lectin receptors such as human dendritic cell inhibitory receptor (DCIR) and blood dendritic cell antigen-2 (BDCA2) interact with the inner mannose residue of complex-type *N*-glycan (Jégouzo et al. 2015; Nagae, Ikeda et al. 2016). These lectins bind either one of the two *N*-glycan branches and show little or no branch preference. On the other hand, DC-SIGN and murine DCIR2 interact with inner mannose residues in branch-specific modes (Feinberg et al. 2001; Nagae et al. 2013). The branch specificity of these C-type lectins is attained at least in part by the ligand conformation, i.e., the shape of the glycan, although the structural basis is still not fully understood (Gabius et al. 2011).

Some other lectins can potentially bind both terminal and inner units of *N*-glycans. Mannose-binding type Jacalin-related lectin (mJRL) is one of the most characterized lectin family and members potentially bind to terminal and/or internal units of *N*-glycans. This family was first found in plants and algae, but recently identified in mammals, drawing attention to structure-function relationships (Hanashima et al. 2015; Kanagawa et al. 2014). mJRLs have a common β-prism fold arranged by three Greek-key motifs (Sankaranarayanan et al. 1996) Figure 1A and B). The first Greek-key motif (β1, β2, β11 and β12) typically contributes to sugar binding via conserved amino acid residues (Figure 1A). mJRLs bind various types of *N*-glycans including high-mannose type and complex-type glycans (Nakamura-Tsuruta et al. 2008). We recently found that Calsepa, one of the mJRLs isolated from *Calystegia sepium*, recognizes the inner mannose residues from two branches of the bisected *N*-glycan (Nagae, Kanagawa et al. 2016). In the crystal, one Calsepa lectin interacts with both α1–3 branch and β-mannose, while the other lectin binds to both α1–6 branch and bisecting GlcNAc. It is noteworthy that two Calsepa lectins sandwich one bisected glycan in a 2:1 manner. The flipped-back conformation of bisected *N*-glycan, which is its preferred conformation, enables the formation of a pseudo-symmetrical sandwich structure. These observations led to our idea that lectins specific for internal sugar residues are useful for analyzing the conformation and interaction of branched *N*-glycans.

Orysata, isolated from *Oryza sativa*, is also a mJRL and shows 30% sequence identity with Calsepa (Figure 1A). Compared with Calsepa, Orysata shows a wider specificity, binding a variety of glycans including high-mannose and complex-type *N*-glycans (Al Atalah et al. 2011). Orysata has potent anti-human immunodeficiency virus and anti-respiratory syncytial virus activity (Al Atalah et al. 2011) and strong insecticidal activity thereby protecting plants from biting-chewing and piercing-sucking insects (Al Atalah et al. 2014). It has been suggested that Orysata can accommodate extended *N*-glycans by recognizing the inner mannose (Al Atalah et al. 2011). Structural analyses of these two lectins toward various *N*-glycans, one with narrow and the other with broad specificity, may provide useful information on carbohydrate recognition mechanisms and conformations of lectin-bound *N*-glycans. We here study Orysata and Calsepa and their binding modes by frontal affinity chromatography (FAC), X-ray crystallography and molecular dynamics (MD) simulations. The roles of the binding and nonbinding branches of bound *N*-glycans in these different mJRL complexes can now be compared.

Materials and methods

Materials

Recombinant Calsepa and Orysata lectins were purchased from Wako Pure Chemicals. The biantennary *N*-glycan unit, Galβ1-4GlcNAcβ1-2Manα1–3[GlcNAcβ1-2Manα1–6]Manα1-*O*-tert-butyldimethylsilyl, was chemically synthesized (glycan I in Figure 1C, (Hanashima et al. 2005)). A glyco-hexapeptide Lys-Val-Ala-Asn-Lys-Thr, in which the asparagine (Asn) side chain is modified with a biantennary *N*-glycan (Galβ1-4GlcNAcβ1-2Manα1–3[Galβ1-4GlcNAcβ1-2Manα1–6]Manβ1-4GlcNAcβ1-4GlcNAcβ1-) from egg yolk was purchased from Fushimi Pharmaceutical Co. Ltd (glycan II in Figure 1C, corresponding to glycan #307 in Figure 2).

Frontal affinity chromatography

Recombinant Calsepa or Orysata was immobilized on *N*-hydroxyl succinimide (NHS)-activated Sepharose 4 Fast Flow at a concentration of 0.3 mg/mL (Calsepa) and 0.5 mg/mL (Orysata) and packed into a column. The column was connected to an automated FAC system (FACII), and equilibrated with 10 mM Tris-HCl, pH 7.6 containing 150 mM NaCl. The column temperature was set to 25°C. A number of pyridylaminated (PA) and *p*-nitrophenyl (*p*NP) glycans derived from *N*-glycans and glycolipids were successively injected into the columns by the auto-sampling system, and elution of each glycan was detected by fluorescence (excitation, 310 nm; emission, 380 nm) or absorbance at 280 nm. Association constants for each glycan were determined as shown in the previous report (Tateno et al. 2007).

Crystallization, data collection, structure determination and refinement

Prior to crystallization experiments, the lectins were concentrated up to 5 mg/mL by Amicon Ultra (molecular weight cut off: 3,000) and mixed with carbohydrate at a final concentration of 1 mM. All crystallization trials were performed by the sitting-drop vapor diffusion method at 293 K. Initial crystallization conditions were searched using Index (Hampton Research).

Diffraction quality crystals of Orysata-biantennary *N*-glycan unit (glycan I) complex and Calsepa-biantennary *N*-glycan (glycan II) complex were obtained under the conditions of 0.1 M bis-tris (pH 5.5), 0.1 M ammonium acetate, 17% (w/v) polyethylene glycol 10,000 (Index #65) and 0.1 M HEPES (pH 7.5), 2.0 M ammonium sulfate (Index #5), respectively. Prior to diffraction experiments, crystals were rapidly frozen by cryo-stream at 95 K. Crystals of Orysata-glycan I complex were frozen by adding 20% (v/v) ethylene glycol as cryoprotectant, whereas crystals of Calsepa-glycan II complex were directly frozen. Our attempts to obtain crystals of an Orysata-glycan II complex and Calsepa-glycan I complex were unsuccessful.

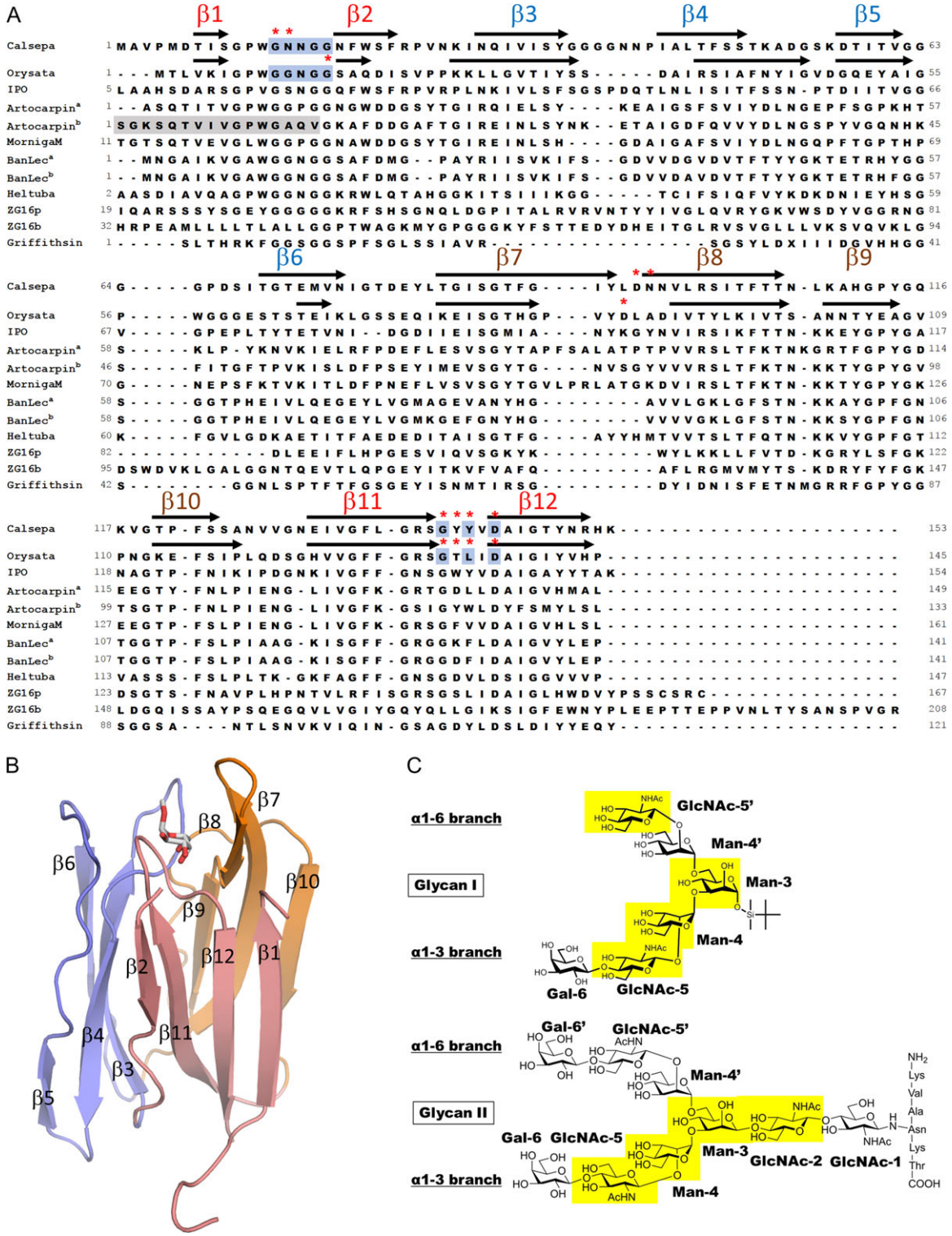

Fig. 1. (**A**) Amino acid sequence alignment of Calsepa, Orysata and other JRL whose 3D structures are available. The alignment was performed with ClustalW. The accession numbers of each sequence are as follows: P93114 (Calsepa), AAB53810 (Orysata), P93193 (IPO), Q7M1T4 (Artocarpin[a]), P18670 and P18671 (α- and β-chains of Artocarpin[b]), Q8LGR3 (MornigaM), O22321 (BanLec[a]), Q8L5H4 (BanLec[b]), Q9ZQY5 (Heltuba), O60844 (ZG16p), Q96DA0 (ZG16b) and P84801 (Griffithsin). The short β-chain of Artocarpin is colored in gray. Conserved GG loop (G-x-x-G-G) and ligand-binding loop (G-x-ϕ-x-D, ϕ: hydrophobic residue) of Orysata and Calsepa are colored in blue. Secondary structures of Calsepa and Orysata are indicated above their sequences. Twelve β-strands are arranged into three Greek-key motifs. β1, β2, β11 and β12 form the first, while β3-β6 and β7-β10 form the second and third motifs, respectively. The three Greek-key motifs are colored red, blue and brown, respectively. Amino acid residues which interact with glycans I and II are indicated with red asterisk above the sequences. (**B**) Representative 3D structure of JRL ((Sankaranarayanan et al. 1996) PDB code: 1JAC) is shown in ribbon model. The three Greek-key motifs are colored as in (**A**). Twelve β strands are labeled. A sugar ligand, α-methyl-galactoside, is depicted in stick model. (**C**) Chemical structures of the ligands in this study. Chemical structures of biantennary hexasaccharide in Orysata complex (upper panel, glycan I) and biantennary glycopeptide in Calsepa complex (lower panel, glycan II). Sugar residues which directly interact with protein are highlighted in yellow. This figure is available in black and white in print and in color at *Glycobiology* online.

Diffraction experiments were performed at BL-5A in the Photon Factory (Tsukuba, Japan). The intensities of diffraction spots were integrated and scaled by HKL2000 (Otwinowsk and Minor 1997). Phase determination of two complexes was performed by the molecular replacement method using program MOLREP (Vagin and Teplyakov 2010). Crystal structures of MornigaM (PDB code: 1XXQ, (Rabijns et al. 2005)) and Calsepa (PDB code: 1OUW, (Bourne et al. 2004)) were used as search models for Orysata and Calsepa complexes, respectively. After phase determination, model building was manually conducted using Coot (Emsley and Cowtan 2004). Refinement was initially performed with Refmac5 (Murshudov et al. 1997) and *phenix.refine* in the Phenix program suite (Adams et al. 2010) for the final models. Stereochemical quality of the models was validated with Molprobity (Chen et al. 2010). Data collection and refinement statistics are summarized in Table S1. All figures were prepared with PyMOL (DeLano Scientific). The atomic coordinates and structure factors of Orysata-glycan I complex and Calsepa-glycan II complex were deposited in the Protein Data Bank under accession codes 5XFH and 5XFI, respectively.

Molecular modeling of Orysata lectin in complex with biantennary glycans

The crystal structure of Orysata reported in this study was used for further MD simulation. Missing residues (Glu60-Thr64 and Glu72) were modeled by the loop modeling approach in MODELLER (version 9.17) using Orysata crystal structure as a template (Eswar et al. 2007). Coordinates of the missing residues from the top scoring (DOPE Score) model were extracted and merged into the crystal structure coordinates. Thus, the coordinates for all the residues except for the missing ones remain those of the crystal structure. The initial structures of the glycan ligands (#301, #302, #304 and #307) were modeled in GLYCAM Web (Woods Group, 2017). Glycans were docked into Orysata manually by superimposing α1–3 branch residues of the glycans on the corresponding residues in the crystal structure. The glycans can potentially bind to Orysata in a second-binding mode, where the α1–6 branch of the glycan occupies the binding site and Man-4 is replaced by Man-4′. The Orysata-glycan complexes in the α1–6 binding mode were generated by superimposing α1–6 branch residues over those of the α1–3 branch in the crystal structure. Additionally, MD of Glycan I and Glycan I-core (Galβ1-4GlcNAcβ1-2Manα1–3[GlcNAcβ1-2Manα1–6]Manβ1-4GlcNAcβ1-4GlcNAc) with Orysata (A) and Oysata (B) was performed to mimic the glycan-binding modes and arrangement of Orysata (A) and (B) in the crystal structures. Similarly, MD of Glycan II and Glycan II-nocore (Galβ1-4GlcNAcβ1-2Manα1–3[Galβ1-4GlcNAcβ1-2Manα1–6]Manβ) was performed to understand the *N*-glycan binding with Calsepa (A) and Calsepa (B) in the crystal structures.

Structures for MD were prepared using the *tleap* module of AmberTools16. The protein and glycans were described by AMBER ff14SB and GLYCAM (version 06j) force fields, respectively. All initial structures were solvated in an octahedral box of TIP4P water molecules extending 12 Å in each direction using *tleap*. All the systems were equilibrated using the multistep equilibration protocol published recently (Nagae et al. 2017). Then, a 300-ns MD simulation at NPT was performed for each case using CUDA implementation of the *pmemd* from AMBER14 (Case et al. 2005). To confirm the binding of glycans in the α1–6 binding mode, this set of MD simulations was extended to 500 ns. In addition to this, an MD simulation of Orysata complexed with glycan I was performed for 300 ns. The temperature was kept constant at 300 K using the Berendsen weak-coupling method. The time constant for heat bath coupling was set to 5 ps. The Particle Mesh Ewald (Darden et al., 1993) was used for calculating electrostatic forces. A cutoff of 9 Å was used for nonbonded interactions, and the SHAKE algorithm (Ryckaert et al. 1977) was used to restrain hydrogen atoms. The MD trajectories were analyzed (hydrogen bond analysis and calculation of dihedral angles) using *cpptraj* utility (Roe and Cheatham, 2013) of AmberTools16. Similarly, 300-ns MD simulations of each of the glycans in water was performed using the same procedure as for protein/glycan complexes. The MD simulation of Orysata/glycans complexes in α1–3 binding mode were repeated for 200 ns to check reproducibility of the binding free energy calculations.

The binding free energies of lectin-glycan complexes were calculated using the molecular mechanics/generalized-Born surface area (MM/GBSA) approach.

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S \approx \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\Delta S$$

A total of 500 snapshots were used to compute enthalpy (ΔH) and 100 for entropy ($T\Delta S$) contributions to the free energy of binding. The molecular mechanics term (ΔE_{MM}) includes energies for bonded (bond, angle and dihedral energies) terms (ΔE_{bonded}) and nonbonded terms ($\Delta E_{\text{nonbonded}}$). The $\Delta E_{\text{nonbonded}}$ includes electrostatic (ΔE_{ele}) and van der Waals (ΔE_{vdw}) energy contributions to binding free energy. Whereas, solvation energy term (ΔG_{solv}) is the sum of the polar contribution to solvation free energy calculated by the Generalized-Born model (ΔG_{GB}), and nonpolar contribution calculated from a linear relation to solvent accessible surface area (ΔG_{SA}). Generalized-Born (GB) continuum model (igb = 7) was used to compute polar contribution of solvation energy (ΔG_{GB}). The *PBRadii* was set to *mbondi*. A probe of size 1.4 Å was used and SURFTEN and SUFROFF parameters were set to 0.005 and 0, respectively. The entropic contributions were calculated on 100 snapshots by harmonic approximation using *nmode* in AmberTools16. The standard error of mean (SEM) of each term was calculated by the standard deviation of each term over all the snapshots divided by total the number of snapshots (500 for enthalpy and 100 for entropic term). Uncertainties in ΔG values were calculated from the propagation of errors in ΔH (∂H) and $T\Delta S$ (∂S) values to ΔG (∂G) as:

$$\partial \mathrm{G} = \sqrt{(\partial H)^2 + (\partial S)^2}$$

To compare the ΔG of each glycan with FAC data, we calculated overall ΔG (ΔG_{avg}) of each glycan as a mean of ΔG for α1–3 and α1–6-binding modes of the glycan assuming α1–3 and α1–6-binding modes occur equally. Glycan conformations in MD were clustered into five clusters, based on root mean square deviation (RMSD) of ring atoms, using *k-means* approach. The RMSD between representative conformations of the most populated clusters in water and lectin-bound MD simulation was calculated. Electrostatic interaction energies of glycans with surrounding water ($\langle V^{el}_{l-s}\rangle_{\text{free}}$) and lectin-bound simulations ($\langle V^{el}_{l-s}\rangle_{\text{bound}}$) were calculated by post processing MD trajectories in *cpptraj module of AMBER*.

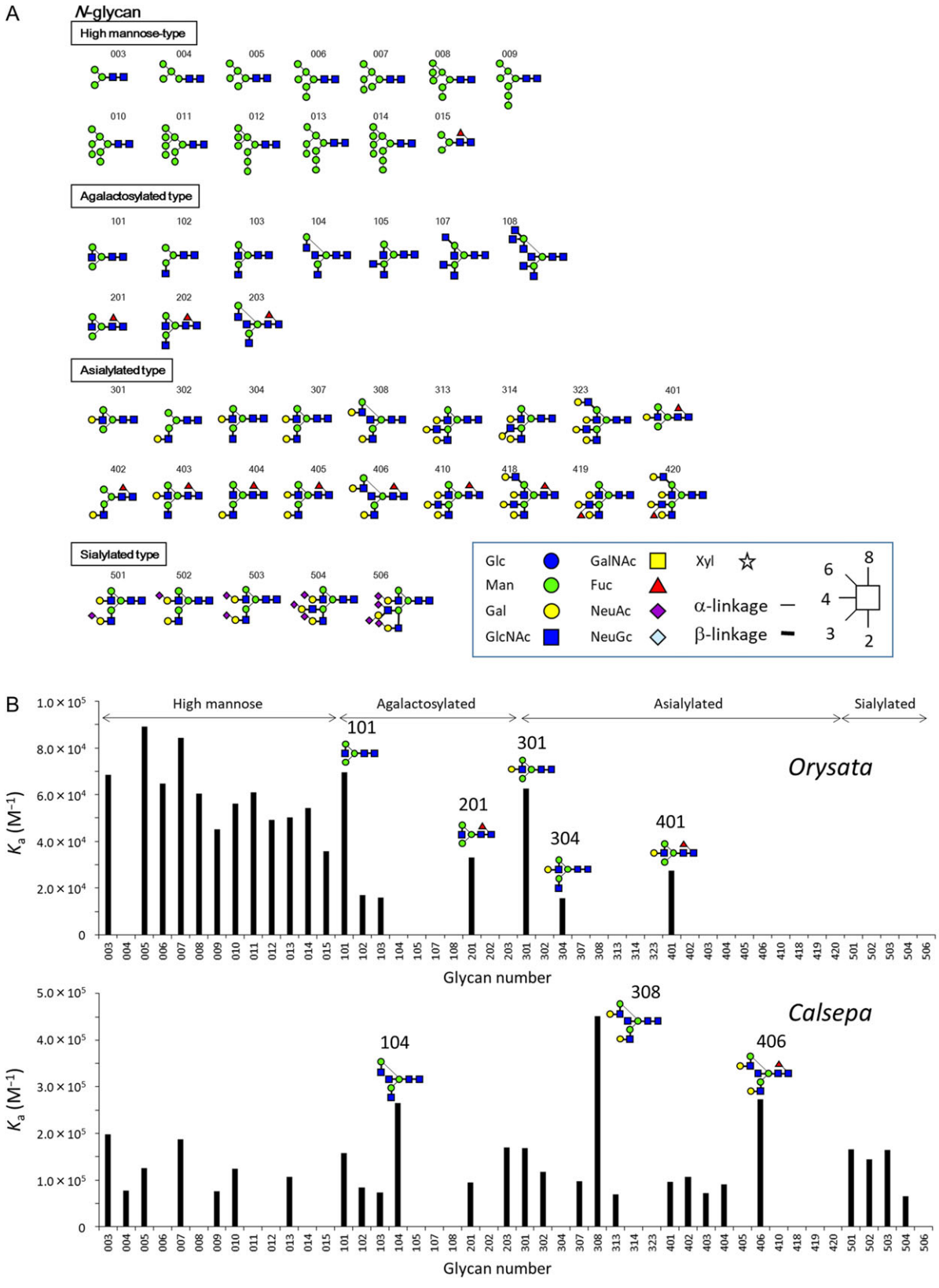

Fig. 2. FAC analysis of Orysata and Calsepa lectins. (**A**) Schematic representation of PA or *p*-nitrophenylated oligosaccharides derived from various types of *N*-linked glycans. Monosaccharide symbols follow the SNFG (symbol nomenclature for glycans) system (Varki et al. 2015). (**B**) Bar graph representation of the association constants (K_a) of Orysata (upper panel) and Calsepa lectin (lower panel) for various *N*-glycans. The numbers at the bottom of the panel correspond to those in (**A**). This figure is available in black and white in print and in color at *Glycobiology* online.

Results

Carbohydrate-binding specificities of Orysata and Calsepa

We here focused on two mJRLs, Calsepa and Orysata. The carbohydrate specificity of Calsepa has previously been investigated by FAC analysis (Nakamura-Tsuruta et al. 2008), but not that of Orysata. Therefore, we performed a FAC analysis for a panel of PA- and *p*NP-linked oligosaccharides using recombinant Orysata and Calsepa under the same conditions (Figure 2A and Supplementary data, Figure S1). The data show that the two lectins bind to *N*-glycans but not to glycolipid-type and other type glycans. Orysata binds various types of *N*-glycans such as high-mannose (#003, #005-#015) and complex-type glycans (#101-#103, #201, #301, #304 and #401) (Figure 2B, upper panel). Interestingly, Orysata binds to all the high-mannose-type glycans except #004, which lacks the α1-3-linked mannose residue attached to core β-mannose. This suggests that the α1-3-linked mannose residue in high-mannose-type glycan is essential for binding to Orysata. Since Orysata can bind several asialo (galacto) and agalacto glycans, the protein seems to recognize the inner part of the complex-type glycans. Intriguingly, Orysata prefers complex-type biantennary glycans with an extension at the α1–6 branch over an extension at the α1–3 branch (#301 vs #302, #401 vs #402). Each branch apparently contributes independently to the Orysata affinity. On the other hand, Calsepa binds several high-mannose-type and complex-type *N*-glycans (Figure 2B, lower panel). Calsepa, but not Orysata, binds #004 suggesting that Calsepa can recognize mannose residue(s) other than the α1–3 branched mannose (Man-4) in high-mannose-type glycans. Among the complex-type glycans tested, Calsepa showed a preference for bisected *N*-glycans which have a bisecting GlcNAc at the core mannose (#104, #308 and #406), consistent with previous reports (Nagae, Kanagawa et al. 2016; Nakamura-Tsuruta et al. 2008).

Crystal structure of Orysata in complex with branched *N*-glycan

We determined the crystal structure of Orysata in complex with a complex-type biantennary glycan unit (glycan I in Figure 1C) at 1.9 Å resolution. As found in typical mJRLs, Orysata has a β-prism fold composed of 12 β-strands (β1-β12) (Figure 3A, left panel). In the asymmetric unit, two Orysata lectins (A and B) sandwich one biantennary *N*-glycan unit in a 2:1 manner. In the electron density map, all six sugar residues are clearly visible between Orysata (A) and Orysata (B) (Figure 3B). This 2:1 sandwich was also seen in a previously reported Calsepa-bisected *N*-glycan complex (right panel in Figure 3A, (Nagae, Kanagawa et al. 2016)). However, the sandwich modes of the two complexes are different. When the two structures were superposed using Orysata (A) and Calsepa (A), the positions of the α1–3 branches superimpose well (Supplementary data, Figure S2). In contrast, the relative positions of Orysata (B) and Calsepa (B) are completely different. Structural superposition of Orysata (B) and Calsepa (B) results in polar angles of the rotation operator ($\phi = 94°$, $\psi = 30°$, $\chi = 125°$).

The defining characteristic of the Orysata-biantennary *N*-glycan complex is the asymmetric ligand recognition by the two lectins. Orysata (A) interacts with Man-3, Man-4 and GlcNAc-5 in the α1–3 branch, whereas Orysata (B) binds to GlcNAc-5′ in the α1–6 branch and GlcNAc-5 in the α1–3 branch (Figure 3C and Table S2). Among the sugar residues directly involved in the lectin interaction, Man-4 (α1–3 branch) mainly fits into the conserved sugar-binding site of Orysata (A), while GlcNAc-5′ (α1–6 branch) interacts with the same binding site of Orysata (B). Three hydroxyl groups, OH3, OH4 and OH6 and ring oxygen atom O5 of Man-4 interact with Orysata (A), and likewise OH3, OH4, OH6 and ring oxygen of GlcNAc-5′ are recognized by Orysata (B). GlcNAc-5 binds to both Orysata (A) and (B) via two sites. The *N*-acetyl oxygen atom (O7) of GlcNAc-5 interacts with the main chain of Gly133 in Orysata (A), while OH3 of GlcNAc-5 makes a hydrogen bond with the side chain of Asp87 in Orysata (B). There is no direct interaction of Gal-6 (α1–3 branch) nor of Man-4′ (α1–6 branch) with Orysata lectins. It seems that the interaction of GalNAc-5′ (α1–6 branch) with Orysata (B) is attained because GlcNAc-5′ is not modified with β1–4 galactose (Gal-6′). The superimposition of the Galβ1-4GlcNAc moiety of the α1–3 branch onto GlcNAc-5′ of the α1–6 branch results in a severe steric clash between Orysata (B) and putative Gal-6′ (Supplementary data, Figure S3). The two Orysata lectins seem most favorably arranged to maximize interaction with the asymmetric branch structure of biantennary glycans.

The global shape of biantennary *N*-glycan is affected by the dihedral angles of the Manα1–6Man linkage and roughly divided into two conformers, extend and folded (back-fold) (Stubbs et al. 1996), which can be further divided into five conformers, extend-a, extend-b, back-fold, half back-fold and tight back-fold (Nishima et al. 2012). The dihedral angles of the Manα1–6Man linkage in the Orysata-glycan I complex are $\phi_{\alpha1-6} = 58°$, $\varphi_{\alpha1-6} = -179°$, and $\omega_{\alpha1-6} = 68°$ (see Table I). This corresponds to an extend-a conformation, which is the prevalent conformation of biantennary glycans in the lectin-free state predicted by MD simulations (Nishima et al. 2012) and is often observed in crystal structures of lectin-glycan complexes (Nagae, Kanagawa et al. 2016).

Crystal structure of Calsepa in complex with biantennary *N*-glycan

We previously solved a crystal structure of Calsepa in complex with a bisected *N*-glycan unit (Nagae, Kanagawa et al. 2016). In this crystal structure, a flipped-back conformation of the bisected glycan unit enables cooperative sandwich interaction with two Calsepa lectins. However, a question still remains as to how Calsepa binds to nonbisected, biantennary *N*-glycan with a chitobiose core. To address this question, we here determined the crystal structure of Calsepa in complex with a nonbisected biantennary *N*-glycan (glycan II in Figure 1C, corresponding to glycan #307 in Figure 2) at 1.65 Å resolution. In the asymmetric unit, two Calsepa lectins (A and B) are found and each Calsepa independently interacts with one glycan molecule (Figure 4A). The electron density from six sugar residues was clearly observed and assigned to the α1–3 branch (Gal-6, GlcNAc-5 and Man-4) and core chitobiose region (Man-3, GlcNAc-2 and GlcNAc-1) (Figure 4B). Additional weak electron density was observed at the 6-position of Man-3, which is assigned to Man-4′ and GlcNAc-5′ from the α1–6 branch. The peptide region of the glycopeptide ligand is fully exposed to solvent and does not seem to contact with either lectins or glycans. Possibly due to this reason, the electron density was indistinct (Supplementary data, Figure S4). The glycan-binding mode is in marked contrast to the previous crystal structure of Calsepa-bisected *N*-glycan sandwich complex (right panel in Figure 3A).

The α1–3 branch (Gal-6, GlcNAc-5 and Man-4) and core chitobiose region (Man-3, GlcNAc-2 and GlcNAc-1), are in a straight

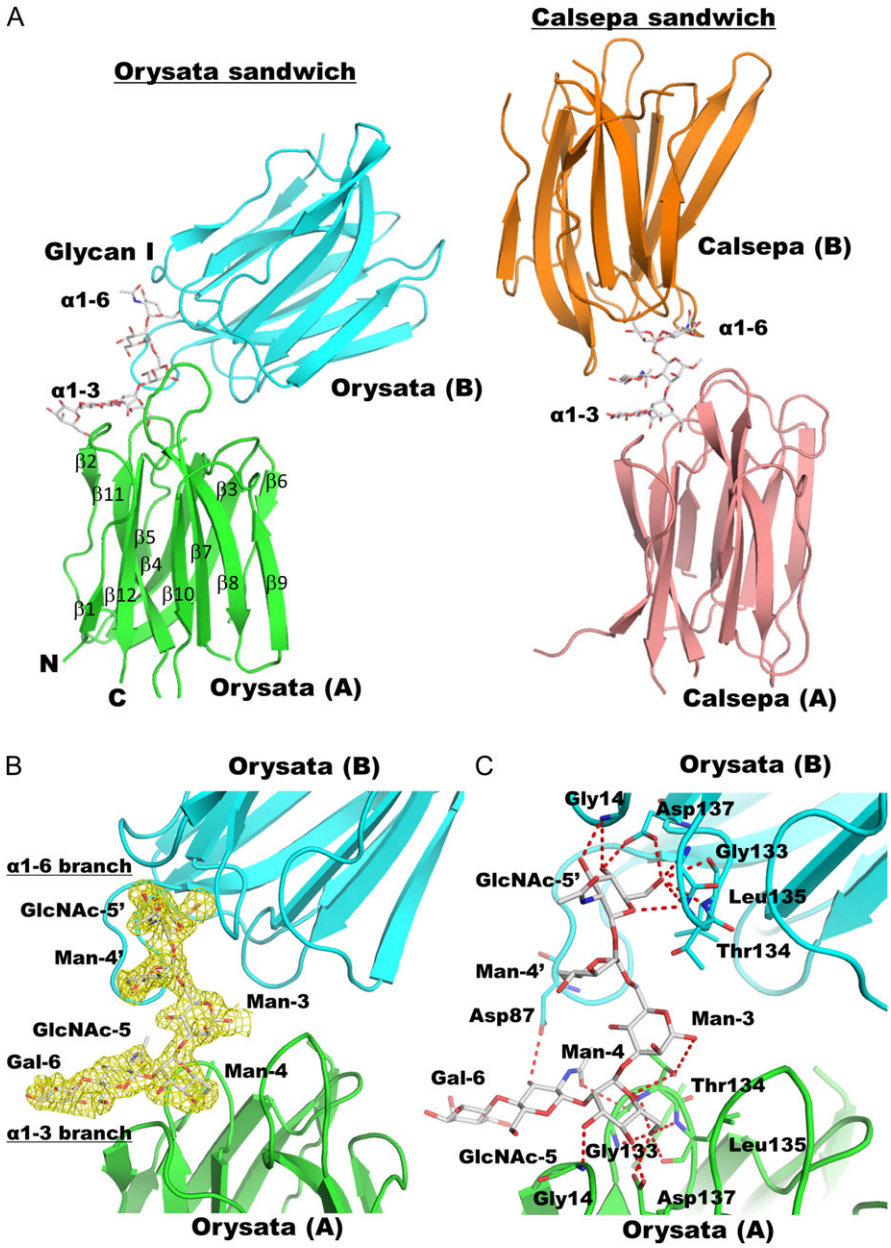

Fig. 3. Crystal structure of Orysata in complex with biantennary *N*-glycan (glycan I). (**A**) Overall structures of Orysata in complex with bianennary glycan (left panel) and Calsepa-bisected glycan complex which was published previously (PDB code: 5AV7, right panel). Two lectins, Orysata (green and cyan) and Calsepa (pink and orange), sandwich one glycan in the two complex structures. Protein and carbohydrate molecules are depicted as ribbon and stick models, respectively. Twelve strands (β1-β12), N-terminus and C-terminus of Orysata (**A**) are indicated. (**B**) Omit map contoured at 2.5 σ level around the sugar-binding site is depicted in yellow mesh. Amino acid residues which interact with carbohydrate are shown in rod models. (**C**) Close up view of the sugar-binding site. Putative hydrogen bonds are shown in red dotted lines. Details of the interaction network are summarized in Supplementary data, Table S2. This figure is available in black and white in print and in color at *Glycobiology* online.

line due to binding to Calsepa (Figure 4C). Among the six sugar residues, Man-4 extensively interacts with Calsepa by fitting into the mannose-binding pocket (Table S3). The OH6 of Man-4 is buried inside the pocket and hydrogen bonds with Calsepa. The OH3 and *N*-acetyl oxygen atom (O7) of GlcNAc-5 interact with the side chain of Asn18. The OH2 of Man-3 hydrogen bonds with the side chain of Asn96 and the nonpolar B-face of Man-3 stacks with the aromatic ring of Tyr142. The linkage oxygen (O4) and *N*-acetyl nitrogen atom (N2) of GlcNAc-2 interact with the side chains of Asp95 and Asn96, respectively. Although GlcNAc-1 and Gal-6 only weakly interact with Calsepa, their electron density is clearly visible (Figure 4B and C). The bisected glycan unit without a chitobiose core exhibits very weak affinity toward Calsepa which precluded estimation of a dissociation constant (Nagae, Kanagawa et al. 2016). It is likely that the direct interaction of GlcNAc-2 with Calsepa contributes to the affinity. Actually, we detected a significant affinity for the corresponding glycan #104 in the FAC analysis (Figure 2B). Collectively, both the chitobiose part and the other arms seem to contribute to attaining strong Calsepa binding. The dihedral angles of the Manα1–6Man linkage in the Calsepa-glycan II complex are $\phi_{\alpha 1-6}$ = 62° and 64°,

Table I. Dihedral angles of biantennary and bisected glycans in three lectin complexes

Lectin Glycan	Orysata Glycan I	Calsepa Glycan II A, B	Bisected glycan[1]
GlcNAcβ1-4GlcNAc (ϕ,φ)	-	(−67, 103), (−69, 105)	-
Manβ1-4GlcNAc (ϕ,φ)	-	(−68, 131), (−69, 130)	-
Manα1–3Man (ϕ,φ)	(67, −130)	(95, −125), (95, −124)	(104, −144)
GlcNAcβ1-2Man (α1–3 branch) (ϕ,φ)	(−108, −118)	(−86, −134), (−87, −133)	(−88, −134)
Galβ1-4GlcNAc (α1-3 branch) (ϕ,φ)	(−80, 127)	(−81, 99), (−83, 96)	-
Manα1–6Man (ϕ,φ,ω)	(63, −179, 63)	(62, 103, 50), (64, 104, 51)	(98, 106, 57)
GlcNAcβ1-2Man (α1-6 branch) (ϕ,φ)	(−77, −8)	(−100, −95), (−101, −96)	(−91, −129)
GlcNAcβ1-4Man (bisecting GlcNAc) (ϕ,φ)	-	-	(−63, 138)
Conformation	Extend-a	Back-fold	Back-fold

The Φ, Ψ and ω angles are defined by atoms O_5–C_1–O'_x–C'_x, C_1–O'_x–C'_x-C'_{x-1} and O'_x–C'_x–C'_{x-1}–C'_{x-2}.
[1]Retrieved from previous report (PDB code: 5AV7, (Nagae, Kanagawa et al. 2016)).

$\varphi_{\alpha 1-6} = -103°$ and $-104°$, $\omega_{\alpha 1-6} = 50°$ and 51° (Table I), categorizing it as a back-fold conformation, which is also observed in the bisected glycan unit-Calsepa complex (Nagae, Kanagawa et al. 2016). Although the α1-6 branch is exposed to solvent and does not directly interact with Calsepa, the branch seems stabilized via an intramolecular hydrogen bond between the *N*-acetyl nitrogen atom of GlcNAc-5′ and the OH3 group of GlcNAc-2 (Figure 4C). The dihedral angles of the α1–3 branch of glycan I (in complex with Orysata) are similar to those of glycan II in complex with Calsepa (Table I).

The most striking difference between the nonbisected and the bisected glycan complexes is the independent interaction mode of the former and the sandwich-like interaction mode of the latter. Despite this, the two protein structures are quite similar (Figure 4D), with a RMSD of corresponding 141 Cα atoms of only 0.44 Å. Furthermore, the relative positions of Man-4 and GlcNAc-5 (α1–3 branch), Man-4′ and GlcNAc-5′ (α1–6 branch) and Man-3 (β-mannose) are almost the same. This suggests that the flipped-back conformation is rather energetically stable, and may prevail without bisecting GlcNAc. The introduction of bisecting GlcNAc may further stabilize the flipped-back structure, enabling the sandwich-like binding mode observed in the crystal. Structural superposition between Calsepa-glycan II and Calsepa-bisected glycan complexes shows that there is no apparent steric clash between chitobiose of glycan II and any part of bisected glycan complex (Supplementary data, Figure S5). Thus, the addition of chitobiose does not seem to affect the flipped-back conformation and sandwich formation.

Structural comparison between Calsepa and Orysata

Structural superposition of Orysata and Calsepa yields a RMSD value for 122 Cα atoms of 2.0 Å. In the sugar-binding sites, the positions of the amino acid residues which interact with glycan are well superimposable (Figure 5A). Gly14 in the GG loop (β1-β2 loop) in Orysata structurally corresponds to Gly17 in Calsepa, and Gly133 and Asp137 in the ligand recognition loop (β11-β12 loop) in Orysata correspond to Gly140 and Asp144 in Calsepa, respectively. In the Calsepa-bisected glycan complex, Asn96 in the ligand-binding loop (β7-β8 loop) is involved in symmetrical bisecting GlcNAc recognition. This residue is replaced with Leu88 in Orysata, which must affect the binding property. Moreover, differences in amino acid sequence and 3D structure of the two mJRLs occur in loop regions of the second Greek-key motif including β3-β6 strands (Figure 5B). The differences here, especially in the β3-β4 and β5-β6 loops, seem to dictate the distinct quaternary arrangements of these two lectins in the crystals.

Molecular modeling of glycan-binding modes in Orysata and Calsepa

The crystal structures of Orysata and Calsepa suggest that 2:1 binding is possible in both lectins. The interaction of α1–3 and α1–6 chains of biantennary glycans by Calsepa/Orysata (A) and (B) suggested a possibility of two-binding modes for biantennary glycans in both lectins. To explore this, we calculated binding energies of glycan I and glycan I-core in complex with Orysata (A) and Orysata (B), where two-binding modes are taken directly from the Orysata crystal structure. The chitobiose core was further connected manually to glycan I in a way that avoided steric clashes with Orysata. The binding affinity of the α1–3 branch of glycan I and I-core in Orysata (A) is more favorable than the affinity of α1–6 branch in Orysata (B) (Table S4). Though the absolute binding energies for Glcyan I-core is positive for both binding modes, the relative change in binding energy shows Orysata prefers the α1–3 branch. A similar trend is seen for α1–3 and α1–6 branches of glycan II and II-nocore binding to Calsepa (A) and Calsepa (B), respectively. These results indicate that both branches of biantennary glycans may bind to lectin, resulting in a 2:1 binding. However, each branch binds with different affinity. It can be surmised that there are two-binding modes possible: (I) a stronger binding mode that involves interaction with Man-3 and the chitobiose core and (II) a weaker one with GlcNAc-5′ in the binding site. However, the weaker binding mode may be so weak that it cannot normally be observed in the face of the stronger binding mode. Glycan II in α1–3 binding mode shows that the chitobiose core makes a favorable enthalpy contribution. However, there is no significant difference in the overall binding affinities of glycan II and II-core in α1–3 or α1–6 binding modes (Table S4). This result excludes the possibility that the presence of a chitobiose core in glycan leads to 1:1 binding in Calsepa. Thus, 1:1 binding in Calsepa-glycan II complex can be attributed to the influence of the peptide attached to glycan II.

MD simulations of Orysata complexed with branched *N*-glycans

The glycan-binding specificity of Calsepa is well explained by the crystal structures of glycan-Calsepa complexes. However, that of Orysata is rather difficult to understand from the crystallographic

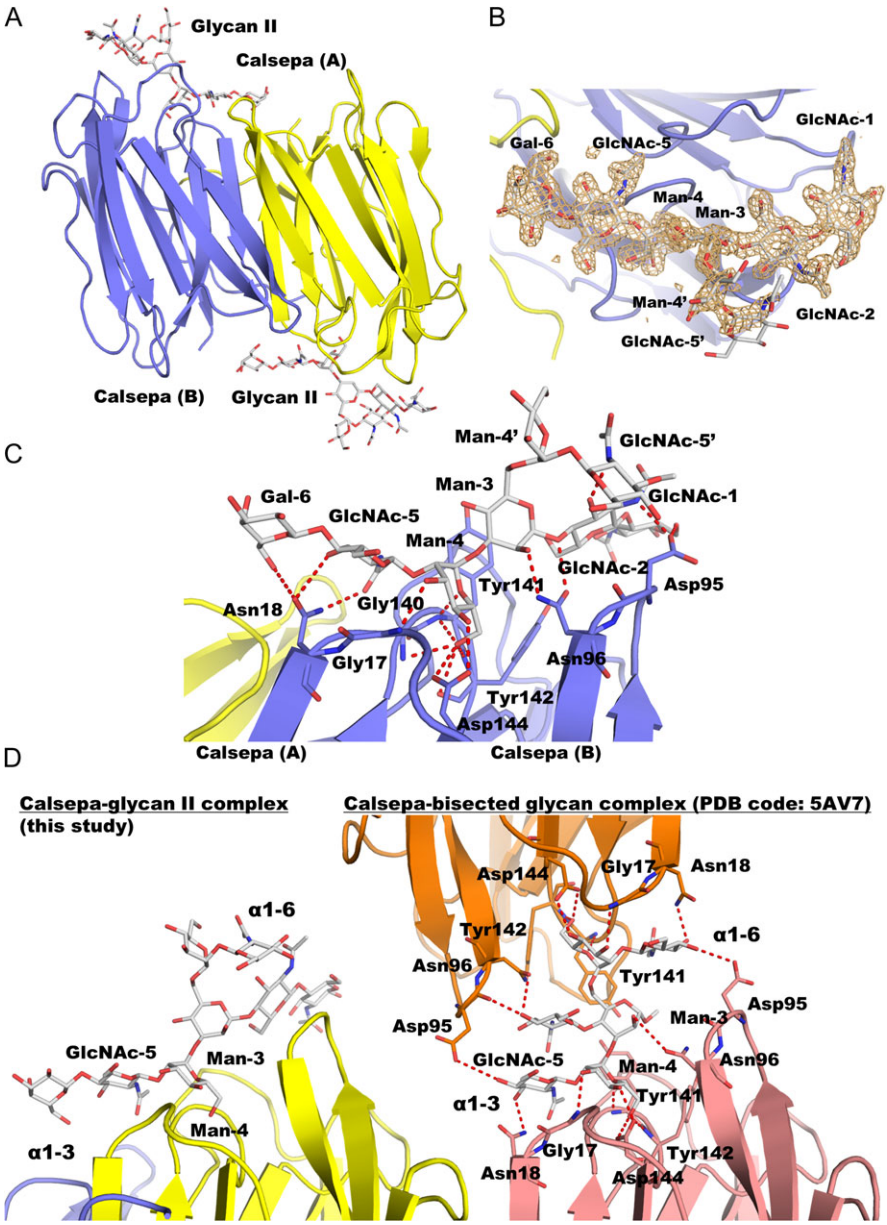

Fig. 4. Crystal structure of Calsepa in complex with biantennary *N*-glycan (glycan II). (**A**) Overall structure of Calsepa in complex with biantennary glycan in the asymmetric unit. Two protein and two sugar molecules in the asymmetric unit are indicated. (**B**) Omit map contoured at 2.5 σ level around the sugar-binding site is depicted in wheat-colored mesh. Sugar residues are labeled. (**C**) Close up view of the sugar-binding site. Putative hydrogen bonds are shown in red dotted lines. Details of the interaction network are summarized in Supplementary data, Table S3. (**D**) Structural comparison between Calsepa-glycan II complex (left panel) and Calsepa-bisected glycan complex (right panel, PDB code: 5AV7). The interaction network in Calsepa-bisected glycan complex is shown in red dotted lines. This figure is available in black and white in print and in color at *Glycobiology* online.

data, especially the preference toward complex-type *N*-glycans with extension at the α1–6 branch (#307 to #302) and α1–3 branch (#301 to #304 to #302) observed in FAC analysis. In order to further elucidate this, we performed MD simulations of glycan I in the presence of Orysata (Figure S6). We built a starting model of Orysata (A) complexed with glycan I, in which Man-4 (Manα1–3 branch) is embedded in the mannose-binding pocket. We assumed that the 2:1 sandwich-like interaction observed in the crystal is unlikely to occur under FAC experimental conditions because the lectins are covalently immobilized to the column resin. The MD simulation of glycan I bound to Orysata shows that the conformation of the Manα1–6 branch in glycan I is flexible (Figure S6). Some of the glycan conformations seen in the MD trajectory align well with that in the crystal structure (Figure S6-B). Glycan I conformation in the crystal structure is also observed in MD of glycan I in solution (33% population) and lectin-bound state (35% population) (Table S5). The RMSD between the representative conformations of the most populated clusters in both states is 1.4 Å. It suggests that the Glycan I conformation seen in the crystal structure is also prevalent in solution (Table S5).

We next performed MD simulations of four branched glycans, binders (#301 and #304) and nonbinders (#302 and #307), both in the free and bound states with Orysata (A) to understand their

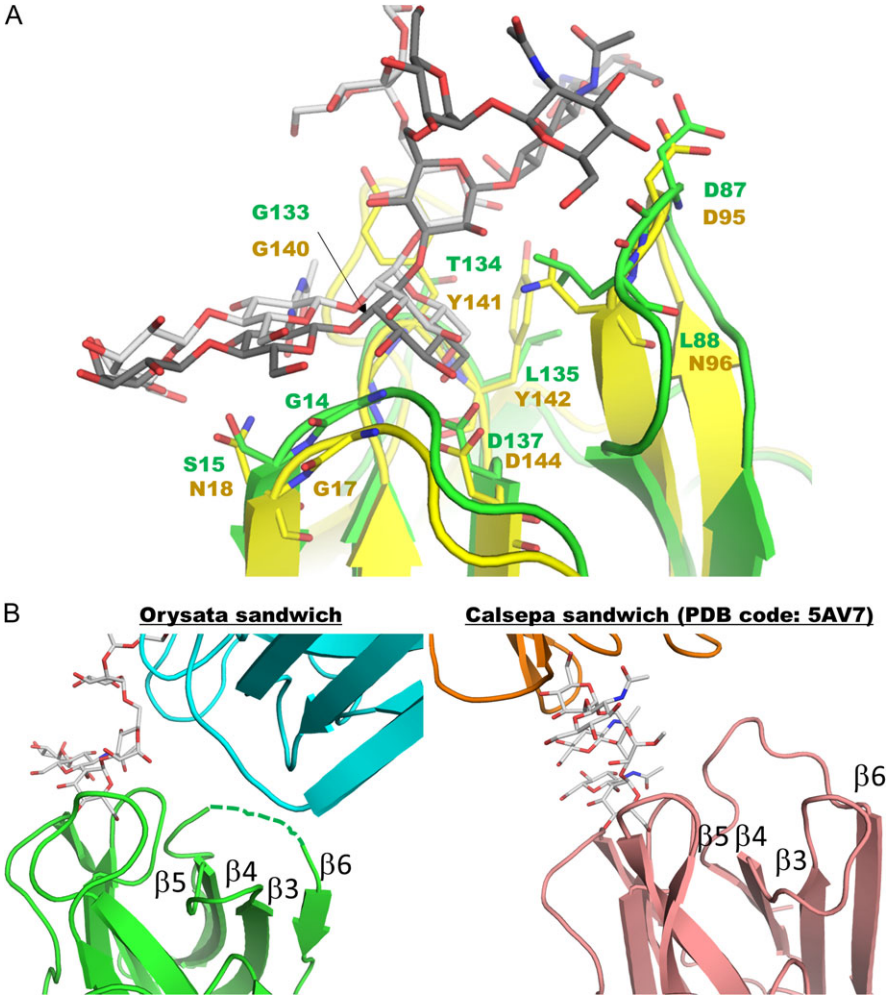

Fig. 5. Structural comparison of Orysata-glycan I, Calsepa-glycan II and Calsepa-bisected glycan complexes. (**A**) Structural superposition of Orysata-glycan I (lectin: green, glycan: white) and Calsepa-glycan II (lectin: yellow, glycan: gray) around sugar-binding sites. Amino acid residues which interact with glycans are shown in rod models and labeled. (**B**) Structural comparison between Orysata-glycan I sandwich (left panel) and Calsepa-bisected glycan sandwich (right panel, PDB code: 5AV7). Four strands of the second Greek-key motif are labeled. This figure is available in black and white in print and in color at *Glycobiology* online.

affinity differences seen in FAC data. In the case of glycan-lectin complexes, Man-4 (α1–3 branch) or Man-4′ (α1–6 branch) was set in the mannose-binding pocket for each glycan (Figure S7). The overall structure of the glycans is represented by the orientation of the three units linked to the core Man-3: α1–3 branch, α1–6 branch and the chitobiose unit. To access the conformational flexibility of the glycans in solution (lectin-free) and in the lectin-bound state, we plotted the dihedral angles of α1–3, α1–6 and Manβ1-4GlcNAc glycosidic linkages (Figure S8-S9). Glycan conformations in solution and in the lectin-bound state do not reflect any particular stable and unique bound state conformation for any of the glycans. However, the most populated conformation of #301, #304 and #307 in the bound state is also prevalent in solution (Table S5 and Figure S10). The overall conformation of #302 in water is also similar to that in the lectin-bound state but the high RMSD (7 Å) is mainly due to flipping of the chitobiose core. However, the decreased flexibility of glycans in the lectin-bound state is apparent (see Figure S8-S9), which can be attributed to the intermolecular interaction of Man-4 and Man-3 with Orysata. Moreover, GlcNAc-1 and GlcNAc-2 form hydrogen bonds with Ser34 and Thr134 in all four cases and this interaction ultimately rigidifies the core chitobiose unit (Table S6). Likewise, Man-4 (α1–3 branch)/Man-4′ (α1–6 branch) in α1–6/α1–3-binding mode forms strong and stable hydrogen bonds with Asp137 in all cases. The nonbinding branch oriented toward the solvent in both α1–3 (the α1–6 of branch of #301, #304 and #307) and α1–6 (the α1–3 branch #302 and #304) binding modes is quite flexible and it attains several conformations including a flip back conformation in #301 (Figure 6A). The back flipping of the α1–6 branch to the chitobiose core is stabilized by inter-branch hydrogen bonds (Figure 6B). Since the Man-4′ alone in #302 can adopt a conformation similar to Man-4′ in back-fold conformation of #301, the flexibility along the α1–6 linkage in α1–3 binding mode cannot be solely attributed to the interactions of GlcNAc-5′ and Gal-6′ with the chitobiose unit or with water (Figure 6C). The high flexibility along the α1–6 glycosidic linkage is not surprising because it contains three variable torsional angles compared to two in other glycosidic linkages.

Hydrogen bond analysis of the MD trajectories suggests that Orysata-glycan interactions are moderately polar in nature, as just a couple of strong hydrogen bonds are observed with 100% occupancy during the MD simulation (Table S6 and Figure S11). This is further corroborated by calculating the changes in the electrostatic

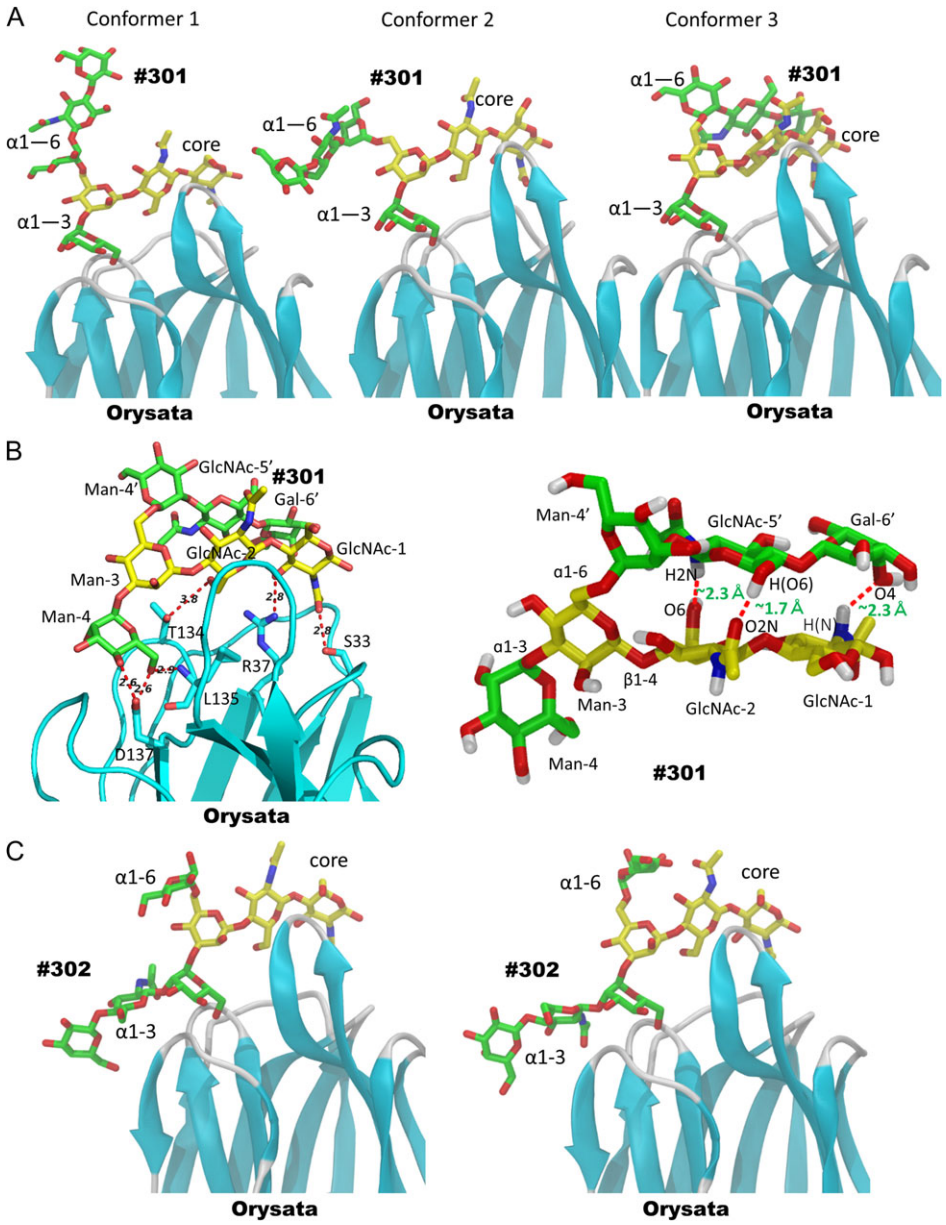

Fig. 6. Conformation of glycans during the α1-3-binding mode MD simulation of Orysata-glycan complexes. (**A**) Three major conformations (Conformers 1, 2 and 3) of the α1-6 branch of #301. (**B**) Flipped-back conformation of #301 showing possible hydrogen bonds between glycan and Orysata (left panel) and between chitobiose and α1-6 branch (right panel). (**C**) Two representative conformations of Man-4′ of #302 observed in MD. Man-4′ (α1-6 branch) is directed toward the α1-3 branch (left) or flipped-back to the core chitobiose (right). This figure is available in black and white in print and in color at *Glycobiology* online.

interaction energies (ΔV^{el}_{l-s}) of the glycans in water and bound state, which vary from −33 to −67 kcal mol^{-1} (Table S7). This range of change in electrostatic interaction energy of glycan is still not as favorable as ΔV^{el}_{l-s} for monosaccharides binding to a bacterial lectin PA-IIL, a typical example of a highly polar interaction in lectin-glycan complexes. (Mishra et al. 2012). The most crucial hydrogen bonds are observed between OH6 and OH4 hydroxyls of Man-4/Man-4′ (in the sugar-binding site) and Asp137 in all four cases (2.6 Å, ~100% occupancy in MD simulation). In addition to this, the OH6 of Man-4/-4′ also forms another hydrogen bond with Leu135 backbone amide, though not with 100% occupancy (Table S6). The chitobiose unit can also form three hydrogen bonds. One is seen between Thr136 and GlcNAc-2 (O6) in the α1–3-binding mode. It has significant occupancy (>50%, $dist_{avg}$ ≤3 Å) for #302, #304, and #307 but exhibits slightly higher $dist_{avg}$ (>3 Å) for #301. In addition to this, and for binders #301 and #304 only, the acetyl group and O3 of GlcNAc-1 make hydrogen bonds with Ser34 and Ser38 in both α1–3 and α1–6-binding modes.

Binding free energies of Orysata-Glycan complexes

For further understanding of the complex-type biantennary glycan recognition by Orysata, we calculated binding energies of all four glycans in both Man-4 (α1–3) and Man-4′ (α1–6) binding modes using the MM/GBSA approach (Table II and Figure 7). In agreement with the FAC data, free-energy calculations show both #301 and

Table II. The binding free energies (±standard error of mean) of four glycans (#301, #304, #302 and #307 in Figure 2A) with Orysata calculated by MM/GBSA (igb = 7) approach are listed in the order of their average ΔG values (ΔG_{avg})

Glycan	branch	ΔE_{ele}	ΔE_{vdw}	ΔG_{polar}	$\Delta G_{nonpolar}$	ΔH	$T\Delta S$	ΔG_{cal}	ΔG_{avg}[1]
#301	α1–3	−107.4 ± 20.8	−42.6 ± 4.9	115.6 ± 17.7	−4.6 ± 0.4	−39.1 ± 7.7	−36.1 ± 4.1	−3.0 ± 0.5	−3.6
	α1–6	−100.4 ± 15.3	−51.9 ± 4.5	112.5 ± 12.6	−5.2 ± 0.3	−45.1 ± 7.2	−41.0 ± 3.1	−4.1 ± 0.4	
#304	α1–3	−87.2 ± 17.6	−47.9 ± 4.9	99.2 ± 15.7	−4.7 ± 0.3	−40.5 ± 6.9	−39.2 ± 4.6	−1.4 ± 0.6	−3.2
	α1–6	−110.4 ± 14.8	−54.3 ± 4.3	122.1 ± 12.5	−5.4 ± 0.2	−48.1 ± 7.0	−43.1 ± 4.1	−5.0 ± 0.5	
#302	α1–3	−86.4 ± 16.7	−52.4 ± 4.0	106.1 ± 12.3	−5.1 ± 0.2	−37.9 ± 7.5	−39.2 ± 2.8	1.3 ± 0.4	−2.3
	α1–6[2]	−103.1 ± 27.0	−38.0 ± 8.2	103.2 ± 23.3	−4.1 ± 0.8	−42.2 ± 11.7	−36.3 ± 6.7	−5.9 ± 0.9	
#307	α1–3	−95.0 ± 19.2	−51.7 ± 4.8	112.6 ± 18.1	−5.1 ± 0.4	−39.3 ± 6.9	−41.1 ± 5.2	1.9 ± 0.6	1.2
	α1–6	−91.1 ± 17.1	−42.2 ± 8.7	101.5 ± 17.4	−4.3 ± 0.7	−36.1 ± 8.0	−36.6 ± 5.1	0.5 ± 0.6	

The ΔE_{ele} and ΔE_{vdw} are the standard molecular mechanics (MM) energy terms from electrostatic and van der Waals interactions. The ΔG_{polar} and $\Delta G_{nonpolar}$ are the polar and nonpolar contribution to the solvation free-energy. The ΔH, $T\Delta S$, ΔG_{cal} denote enthalpy, entropy and binding free energy terms. All the values are in kcal·mol^{-1}. Values other than ΔG_{cal} values are stated with ± standard deviation over all the snapshots.

[1]ΔG_{avg} is calculated as the mean of ΔG_{cal} for α1-3 and α1-6 binding modes, assuming that the two-binding modes occur in equal amounts.

[2]High standard deviation values are due to several structurally different conformations of #302 in α1-6 binding mode.

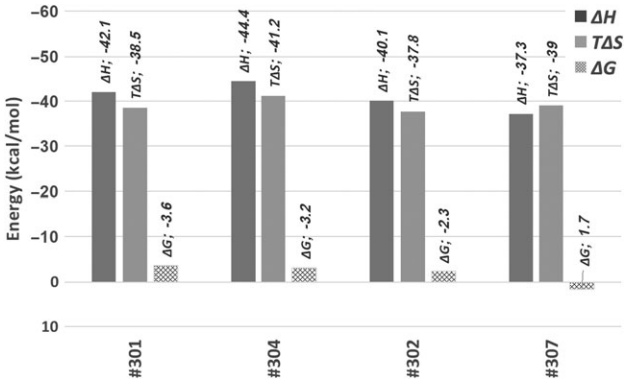

Fig. 7. Plot of average enthalpy (ΔH), entropy ($T\Delta S$) and total binding free energy (ΔG) for the α1-3 and α1-6 branch of glycan #301, #304, #302 and #307. All the values are shown in kcal·mol^{-1}.

#304 bind to Orysata with reasonable affinity, whereas #302 and 307 interact weakly or not at all (last column, Table II). The two possible binding modes for #301 and #304 suggest mannose residues (Man-4 and 4′) of both α1–3 and α1–6 branches bind well, especially through hydrogen bonds with Asp137, and contribute significantly to the binding affinity (Figure S7 and Table S6). Glycan #302 and #307 did not dissociate from the lectin-binding site during the 300-ns MD simulation (α1–3 binding mode) and the 500-ns MD simulation (in α1–6 binding mode), but the weak or zero interaction can be deduced from the binding free energies. While MM/GBSA calculations may not be quantitatively all that accurate, they are qualitatively in good agreement with the FAC data. At present MM/GBSA is the only viable way of obtaining enthalpy and entropic contributions to absolute free energies of binding of large glycans in two different binding modes (considering computational time and accuracy for large flexible ligands).

Binding of glycans can be favored by entropic gain arising from the displacement of water molecules by nonbinding but proximal residues (Navarra et al. 2017), and the methodology used here can help to rationalize enthalpy (using MM/GBSA) and entropic (by normal mode analysis) compensation contributions in such protein/glycan complexes. Glycan-binding preference and extension of α1–3 branch of complex-type *N*-glycans can be explained by the individual contribution of each monosaccharide in binding seen in the free-energy data (Figure 7). The introduction of GlcNAc-5 in the α1–3 branch (#301 to #304) led to a small gain in enthalpy, but the further addition of Gal-6 (#304 to #307) reverses the gain in both binding modes. Despite the gain in enthalpy from #301 to #304, there is an entropic cost that makes #304 the weaker binder. An introduction of residues in the binding branch (#301→304→307 in α1–3-binding mode and #302→#301, #302→#304, #302→#307 in α1–6-binding mode) resulted in more entropic cost. Thus, an extension of the binding branch results in a higher entropic penalty. Similarly, the extension of the nonbinding branch in α1–3 binding mode (#302→#301, #302→#304, #302 to #307) leads to more entropic cost. However, in α1–6 binding mode, extension of α1–3 branch shows complex behavior. $T\Delta\Delta S$ for #301 to #304 shows entropy loss but further extension to #307 shows a substantial gain in entropy. This partially accords with the observation that the entropic contribution due to the introduction of nonbinding monosaccharide moieties in the glycan ligand is complex and crucial to the understanding of protein-glycan interactions (Navarra et al. 2017). The absolute binding energies of *N*-glycans can be reproduced within the limits of computational error in MM/GBSA calculations (Table S8). We would, however, like to point out that due to the large size of the *N*-glycans and difficulty in sampling their conformation flexibility along the branches, binding energies from several independent simulations may differ slightly, but qualitatively the same trends are expected.

Discussion

A fundamental question is how lectins acquire their own ligand-binding specificities. Several lectins show binding preference for specific branches of complex-type *N*-glycan. For instance, a legume lectin PELa shows binding preference for mannose-terminated α1–6 branch and longer α1–3 branch (Benevides et al. 2012). The specificity of PELa is related to the extended binding site and the conformational constrains on the glycan ligand. Another example is a C-type lectin receptor murine DCIR2, which has a strict specificity toward bisected glycan with GlcNAc-terminated α1–3 branch (Nagae et al. 2013). Murine DCIR2 directly interacts with both α1–3 branch and bisecting GlcNAc, hereby the specificity is attained. Alternatively, one concept proposes that the specific glycan-lectin interaction is affected principally by the shape of the glycan, and not only through direct binding (Gabius et al. 2011). Subsequently, however, there have been few examples supporting the concept. We considered that mJRLs may be good models to understand this problem. Although mJRLs possess narrow sugar-binding sites, they have a wide

specificity toward *N*-glycans (Nakamura-Tsuruta et al. 2008). The 3D structural basis for this is not well understood. Here, we analyzed ligand recognition of two mJRLs Orysata and Calsepa toward larger biantennary *N*-glycans by FAC analysis, protein crystallography and MD simulations.

FAC analysis clearly shows that Orysata and Calsepa bind to different sets of glycans. Orysata binds to high-mannose-type glycans except for glycans lacking the Man-4 residue (#004). The crystal structure of the glycan I-Orysata complex shows that Man-4 fits into the sugar-binding pocket. These observations suggest that Man-4 is the main epitope recognized by Orysata. However, Orysata shows a binding preference for complex-type glycans with an extension at the α1–6 branch. This unique ligand specificity will be discussed below from X-ray crystallographic data and MD simulation results.

In the crystal structure, two Orysata lectins bind to one glycan I molecule in an asymmetric sandwich interaction mode: i.e., one Orysata binds Man-4 and the other GlcNAc-5′ of the branched glycan. This is in contrast with the symmetrical binding mode seen in a Calsepa-bisected *N*-glycan complex (Nagae, Kanagawa et al. 2016) in which two Calsepa lectins interact with Man-4 or Man-4′ of the bisected glycan. The binding of GlcNAc-5′ to Orysata is blocked by the Gal-6′ modification (Figure S1). However, FAC analysis indicates Gal-6′ is not inhibitory (#101 vs #301 or #201 vs #401), suggesting that GlcNAc-5′ binding is not significant.

Glycan I holds an extended conformation in the crystal structure (Figure 3A and Table I) and MD simulation of the Orysata-glycan I complex suggests that it occurs in significant amounts in solution (Figure S6). The extend conformation also appears in lectin-glycan complexes such as PAL lectin (PDB code: 1LGB, 2 ARX and 3ZYR), DC-SIGN (PDB code: 1K9I), and galectin-1 (PDB code: 1SLC) (Nagae, Kanagawa et al. 2016), suggesting that the extended conformation is one of the more favored conformations of biantennary *N*-glycans.

It has been reported that bisecting GlcNAc stabilizes a flipped-back conformation of biantennary glycan (Andr et al. 2007; Nishima et al. 2012), and the conformation is likely targeted by Calsepa (Nagae, Kanagawa et al. 2016). In the crystal structure of Calsepa-nonbisected glycan (glycan II, corresponding to glycan #307 in FAC analysis) complex, the α1–6 branch unexpectedly adopts a back-fold conformation despite being exposed to solvent without interacting with the second Calsepa lectin (Figure 4D). Thus, the back-fold conformation seems to occur in solution even in the absence of bisecting GlcNAc. The affinity of glycan #307 is weaker than that of glycan #308 which has additional bisecting GlcNAc (Figure 2B). This difference is supposed to be derived from the direct interaction between Calsepa and bisecting GlcNAc (Nagae, Kanagawa et al. 2016). Or else, the stability of flipped-back conformation may also contribute to the affinity.

In conclusion, we get insights into the divergent binding specificities of two mJRLs, Orysata and Calsepa by experimental and computational approaches. Especially, MD simulations and calculation of binding free energies provide insights into the multiple binding modes: α1–3-binding mode and α1–6-binding mode. Both binding modes contribute to the ligand-binding affinity as supported by the fact that the average binding free energies for both binding modes are qualitatively in agreement with the FAC data. However, the molecular basis of glycan recognition by Orysata is not simply explained by enthalpy-entropy compensation and further analysis is required especially in the interpretation of the entropic contribution. What we can conclude from our analysis is that the nonbinding region of a branched glycan ligand potentially affects its affinity.

Supplementary data

Supplementary data are available at *Glycobiology* online.

Funding

This work was supported in part by Grants-in-Aid for Young Scientists (B) [15K18496 to M.N.], Scientific Research (C) [17K07303 to M.N. and 25460054 to Y.Y.], and Scientific Research (B) [16H04758 to Y.Y.] from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT). This work was also supported by the Mizutani Science Foundation [to Y.Y.].

Acknowledgements

We thank Shoko Wanibe (Wako Pure Chemical Industries, Ltd.) for discussions on Orysata and Noriko Tanaka (RIKEN) for secretarial assistance. This research used the computational resources of RIKEN Advanced Center for Computing and Communication provided under project ID: Q17318. SKM is a research fellow of the Tokyo Biochemical Research Foundation.

Conflict of interest statement

None declared.

References

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW et al. 2010. *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 66:213–221.

Al Atalah B, Fouquaert E, Vanderschaeghe D, Proost P, Balzarini J, Smith DF, Rougé P, Lasanajak Y, Callewaert N, Van Damme EJ. 2011. Expression analysis of the nucleocytoplasmic lectin 'Orysata' from rice in *Pichia pastoris*. *FEBS J*. 278:2064–2079.

Al Atalah B, Smagghe G, Van Damme EJ. 2014. Orysata, a jacalin-related lectin from rice, could protect plants against biting-chewing and piercing-sucking insects. *Plant Sci*. 221–222:21–28.

André S, Kožár T, Schuberth R, Unverzagt C, Kojima S, Gabius HJ. 2007. Substitutions in the *N*-glycan core as regulators of biorecognition: the case of core-fucose and bisecting GlcNAc moieties. *Biochemistry*. 46: 6984–6995.

Benevides RG, Ganne G, Simões Rda C, Schubert V, Niemietz M, Unverzagt C, Chazalet V, Breton C, Varrot A, Cavada BS et al. 2012. A lectin from *Platypodium elegans* with unusual specificity and affinity for asymmetric complex *N*-glycans. *J Biol Chem*. 287:26352–26364.

Bourne Y, Roig-Zamboni V, Barre A, Peumans WJ, Astoul CH, Van Damme EJ, Rougé P. 2004. The crystal structure of the *Calystegia sepium* agglutinin reveals a novel quaternary arrangement of lectin subunits with a β-prism fold. *J Biol Chem*. 279:527–533.

Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ. 2005. The Amber biomolecular simulation programs. *J Comput Chem*. 26:1668–1688.

Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. *MolProbity*: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. 66:12–21.

Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: An *N* log(*N*) method for Ewald sums in large systems. *J Chem Phys*. 98:10089–10092.

Drickamer K, Taylor ME. 2015. Recent insights into structures and functions of C-type lectins in the immune system. *Curr Opin Struct Biol*. 34:26–34.

Emsley P, Cowtan K. 2004. *Coot*: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. 60:2126–2132.

Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A. 2007. Comparative protein structure modeling using MODELLER. *Curr Protoc Protein Sci*. 50:2.9:1–31.

Feinberg H, Mitchell DA, Drickamer K, Weis WI. 2001. Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science*. 294:2163–2166.

Gabius HJ, André S, Jiménez-Barbero J, Romero A, Solís D. 2011. From lectin structure to functional glycomics: Principles of the sugar code. *Trends Biochem Sci*. 36:298–313.

Hanashima S, Götze S, Liu Y, Ikeda A, Kojima-Aikawa K, Taniguchi N, Varón Silva D, Feizi T, Seeberger PH, Yamaguchi Y. 2015. Defining the interaction of human soluble lectin ZG16p and mycobacterial phosphatidylinositol mannosides. *ChemBioChem*. 16:1502–1511.

Hanashima S, Manabe S, Ito Y. 2005. Divergent synthesis of sialylated glycan chains: combined use of polymer support, resin capture-release, and chemoenzymatic strategies. *Angew Chem Int Ed Engl*. 44:4218–4224.

Jégouzo SA, Feinberg H, Dungarwalla T, Drickamer K, Weis WI, Taylor ME. 2015. A novel mechanism for binding of galactose-terminated glycans by the C-type carbohydrate recognition domain in blood dendritic cell antigen 2. *J Biol Chem*. 290:16759–16771.

Kanagawa M, Liu Y, Hanashima S, Ikeda A, Chai W, Nakano Y, Kojima-Aikawa K, Feizi T, Yamaguchi Y. 2014. Structural basis for multiple sugar recognition of Jacalin-related human ZG16p lectin. *J Biol Chem*. 289:16954–16965.

Mishra SK, Sund J, Åqvist J, Koča J. 2012. Computational prediction of monosaccharide binding free energies to lectins with linear interaction energy models. *J Comput Chem*. 33:2340–2350.

Murshudov GN, Vagin AA, Dodson EJ. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*. 53:240–255.

Nagae M, Ikeda A, Hanashima S, Kojima T, Matsumoto N, Yamamoto K, Yamaguchi Y. 2016. Crystal structure of human dendritic cell inhibitory receptor (DCIR) C-type lectin domain reveals the binding mode with *N*-glycan. *FEBS Lett*. 590:1280–1288.

Nagae M, Kanagawa M, Morita-Matsumoto K, Hanashima S, Kizuka Y, Taniguchi N, Yamaguchi Y. 2016. Atomic visualization of a flipped-back conformation of bisected glycans bound to specific lectins. *Sci Rep*. 6:22973.

Nagae M, Mishra SK, Neyazaki M, Oi R, Ikeda A, Matsugaki N, Akashi S, Manya H, Mizuno M, Yagi H et al. 2017. 3D structural analysis of protein *O*-mannosyl kinase, POMK, a causative gene product of dystroglycanopathy. *Genes Cells*. 22:348–359.

Nagae M, Yamaguchi Y. 2014. Three-dimensional structural aspects of protein-polysaccharide interactions. *Int J Mol Sci*. 15:3768–3783.

Nagae M, Yamaguchi Y. 2015. Sugar recognition and protein-protein interaction of mammalian lectins conferring diverse functions. *Curr Opin Struct Biol*. 34:108–115.

Nagae M, Yamanaka K, Hanashima S, Ikeda A, Morita-Matsumoto K, Satoh T, Matsumoto N, Yamamoto K, Yamaguchi Y. 2013. Recognition of bisecting *N*-acetylglucosamine: structural basis for asymmetric interaction with the mouse lectin dendritic cell inhibitory receptor 2. *J Biol Chem*. 288:33598–33610.

Nakamura-Tsuruta S, Uchiyama N, Peumans WJ, Van Damme EJ, Totani K, Ito Y, Hirabayashi J. 2008. Analysis of the sugar-binding specificity of mannose-binding-type Jacalin-related lectins by frontal affinity chromatography—an approach to functional classification. *FEBS J*. 275:1227–1239.

Navarra G, Zihlmann P, Jakob RP, Stangier K, Preston RC, Rabbani S, Smiesko M, Wagner B, Maier T, Ernst B. 2017. Carbohydrate-lectin interactions: an unexpected contribution to affinity. *ChemBioChem*. 18:539–544.

Nishima W, Miyashita N, Yamaguchi Y, Sugita Y, Re S. 2012. Effect of bisecting GlcNAc and core fucosylation on conformational properties of biantennary complex-type N-glycans in solution. *J Phys Chem B*. 116:8504–8512.

Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 276:307–326.

Rabijns A, Barre A, Van Damme EJ, Peumans WJ, De Ranter CJ, Rougé P. 2005. Structural analysis of the jacalin-related lectin MornigaM from the black mulberry (*Morus nigra*) in complex with mannose. *FEBS J*. 272:3725–3732.

Roe DR, Cheatham TE 3rd. 2013. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput*. 9:3084–3095.

Ryckaert JP, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J Comput Phys*. 23:327–341.

Sankaranarayanan R, Sekar K, Banerjee R, Sharma V, Surolia A, 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.

Stubbs HJ, Lih JJ, Gustafson TL, Rice KG. 1996. Influence of core fucosylation on the flexibility of a biantennary N-linked oligosaccharide. *Biochemistry*. 35:937–947.

Tateno H, Nakamura-Tsuruta S, Hirabayashi J. 2007. Frontal affinity chromatography: sugar-protein interactions. *Nat Protoc*. 2:2529–2537.

Vagin A, Teplyakov A. 2010. Molecular replacement with *MOLREP*. *Acta Crystallogr D Biol Crystallogr*. 66:22–25.

Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, Stanley P, Hart G, Darvill A, Kinoshita T et al. 2015. Symbol nomenclature for graphical representations of glycans. *Glycobiology*. 25:1323–1324.

Woods Group. 2005–2017. *GLYCAM Web*. Athens, GA: Complex Carbohydrate Research Center, University of Georgia. http://glycam.org.