京都大学学術情報リポジトリ
KURENAI 紅
Kyoto University Research Information Repository

京都大学
KYOTO UNIVERSITY

Title	Crystal structure of a symbiosis-related lectin from octocoral.
Author(s)	Kita, Akiko; Jimbo, Mitsuru; Sakai, Ryuichi; Morimoto, Yukio; Miki, Kunio
Citation	Glycobiology (2015), 25(9): 1016-1023
Issue Date	2015-09
URL	http://hdl.handle.net/2433/203163
Right	This is a pre-copyedited, author-produced PDF of an article accepted for publication in 'Glycobiology' following peer review. The version of record [Glycobiology (2015) 25 (9): 1016-1023. doi: 10.1093/glycob/cwv033. First published online: May 28, 2015] is available online at: http://glycob.oxfordjournals.org/content/25/9/1016; The full-text file will be made open to the public on May 28, 2016 in accordance with publisher's 'Terms and Conditions for Self-Archiving'.
Type	Journal Article
Textversion	author

Kyoto University

Crystal structure of a symbiosis-related lectin from octocoral

Akiko Kita[1,#], Mitsuru Jimbo[2], Ryuichi Sakai[3], Yukio Morimoto[1], and Kunio Miki[4,#]

[1]Research Reactor Institute, Kyoto University, Kumatori, Sennan, Osaka 590-0494, Japan

[2]School of Marine Biosciences, Kitasato University, 1-15-1, Kitasato, Minami, Sagamihara, Kanagawa 252-0373, Japan

[3]Faculty of Fisheries Sciences, Hokkaido University, Hakodate, Hokkaido 041-8611, Japan

[4]Department of Chemistry, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan

[#]To whom correspondence should be addressed: Akiko Kita (Tel: +81-(72)-451-2379; Fax: +81-(72)-451-2635; E-mail: kita@rri.kyoto-u.ac.jp) or Kunio Miki (Tel: +81-(75)-753-4029; Fax: +81-(75)-753-4032; E-mail: miki@kuchem.kyoto-u.ac.jp)

Proofs and reprints should be addressed to: Akiko Kita, Research Reactor Institute, Kyoto University, Kumatori, Sennan, Osaka 590-0494, Japan.

Running title: Crystal structure of coral symbiosis-related lectin, SLL-2

Key words: coral-algae symbiosis/coral lectin/SLL-2/crystal structure/signal transduction

Abstract

D-Galactose-binding lectin from the octocoral, *Sinularia lochmodes* (SLL-2) distributes densely on the cell surface of microalgae, *Symbiodinium* sp., an endosymbiotic dinoflagellate of the coral, and is also shown to be a chemical cue that transforms dinoflagellate into a non-motile (coccoid) symbiotic state. SLL-2 binds with high affinity to the Forssman antigen, and the presence of Forssman antigen-like sugar on the surface of *Symbiodinium* CS-156 cells was previously confirmed. Here, we report the crystal structures of SLL-2 and its GalNAc complex as the first crystal structures of a lectin involved in the symbiosis between coral and dinoflagellate. *N*-linked sugar chains and a galactose derivative binding site common to H-type lectins were observed in each monomer of the hexameric SLL-2 crystal structure. In addition, unique sugar binding site-like regions were identified at the top and bottom of the hexameric SLL-2 structure. These structural features suggest a possible binding mode between SLL-2 and Forssman antigen-like pentasaccharide.

Introduction

Most tropical corals harbor endosymbiotic dinoflagellates genus *Symbiodinium*, collectively called zooxanthellae. The symbiosis between the corals and zooxanthellae is of particular importance in lives of corals, as the corals relay their nutrient supplies on photosynthetic products of zooxanthellae. Prolonged loss of zooxanthellae causes coral to lose its coloration and the phenomenon was called bleaching. The loss of a healthy symbiotic relationship, triggered by various environmental factors, often results in mass mortality of corals (Baker et al. 2008). It has long been elusive, however, how the corals and the dinoflagellates establish the symbiosis. Because fertilized eggs of corals are mostly symbiont-free, the algal symbiont should be acquired selectively from the adjacent environment during the larval stage (Baker et al. 2004; Koike et al. 2004; Lewis & Coffroth 2004; Sampayo et al. 2008; Adams et al. 2009), and once the algae were acquired in the tissue of coral, healthy reproduction of cells should be maintained. We previously identified a galactosyl glycan binding lectin from a coral *Sinularia lochmodes*, SLL-2 (Jimbo et al. 2000; Jimbo et al. 2005). SLL-2 had an

N-glycosylation site. Three cDNAs closely related to one another were obtained, and the deduced amino acid sequences have a similarity to the H-type lectin family that includes discoidins I and II, slime mold lectins, and *Helix pomatia* agglutinin (HPA) (Poole et al. 1981; Fukuzawa & Ochiai 1996; Sanchez et al. 2006).

Because the lectin is densely localized on the cell surface of *Symbiodinium* within the coral tissue, we assumed some physiological activity of the lectin to the cells (Jimbo et al. 2000). This idea was examined by using various culturable *Symbiodinium* strains and other dinoflagellate cells (Koike et al. 2004). Interestingly, the lectin affected the algal cells in various ways depending on the species or even strains of the algae (Koike et al. 2004). In culture, free living *Symbiodinium* cells change their morphology in diel cycles; i.e.: the cells are in flagellated form and "swim" with circular movement in day time, but the cells transform into an immobilized coccoid form with the loss of flagella at night. The symbiotic forms of *Symbiodinium* cells in corals are "arrested" into the non-motile (coccoid) form, and no diel cycles appear. We took advantage of this unique phenomenon and assessed the effect of SLL-2 on the diel cycle and the cell division of *Symbiodinium* cells *in vitro*. SLL-2 effectively suppressed the diel cycle of

Symbiodinium strain CS-156 with no adverse effects on the cell division. In other strains or species, however, various results were observed, ranging from acute toxicity (cell burst in dinoflagellates *Gymnodinium catenatum* and *Prorocentrum micans*), growth suppression (*Symbiodinium* strain P083-2), cell aggregation to no response (Koike et al. 2004). These results indicated that the binding of the lectin on the cell surface of the microalgae may trigger significant physiological changes in the microalgal cells. The mechanisms underlying these activities are not clear, but galactosyl glycans on the cell surface are the most plausible candidates as the cellular target of the lectin.

We recently found that SLL-2 has the high affinity to the Forssman antigen [*N*-acetylgalactosamine(GalNAc)α1-3GalNAcβ1-3Galα1-4Galβ1-4Glc-ceramide] by frontal affinity chromatography (Jimbo et al. 2013). We also found that proteins that bind to the Forssman antigen, such as HPA and the anti-Forssman glycosphingolipid antibody, bound to CS-156 cells and exhibited physiological changes against the cells. Interestingly, though the effects of HPA were similar to those of SLL-2, those of the anti-Forssman glycosphingolipid antibody were rather suppressive (they reduced the

rate of cell division). Moreover, another Forssman antigen-binding lectin, *Dolichos biflorus* agglutinin (DBA), did not affect the cells. These results suggested that the Forssman antigen-like glycolipid on the surface of the algae is a putative site of lectin binding. These observations also suggested that the structures of the proteins including sugar binding sites possibly control the physiological activity of the lectin against the algal cells.

To facilitate further understanding of the mechanism of the physiological action of the lectin to the symbiotic algae, it is crucial to determine the protein structure. Here, we report the crystal structures of SLL-2 and its complex with GalNAc (SLL-2G) as the first symbiosis-related lectin, and discuss a possible binding mode between SLL-2 and Forssman antigen pentasaccharide chain based on these structures.

Results and discussion

Overall structure of SLL-2 protein

The monomer of SLL-2 is composed of a six-stranded antiparallel β-sandwich consisting of a pair of three-stranded β-sheets (β1, β6, β3, and β2, β5, β4), and extended one β-strand (βN) (Figure 1A). Two intramolecular disulfide bonds are found in the monomer of SLL-2: one between Cys8 and Cys93, and the other between Cys17 and Cys21 (Figure 1A).

Three monomers of SLL-2 are located around a non-crystallographic three-fold axis to form a tight trimeric assembly (Figure 1B). The strands β3 and β4 from the adjacent molecule extend the anti-parallel β-sheet interactions across the interface of the two molecules, resulting in a large bending cleft composed of six continuous stranded, β2, β5, β4, and β3', β6', β1' (Figure 1B, primes denote the adjacent molecule). Based on the calculation of potential on solvent accessible surface (Baker et al. 2001; Dolinsky et al. 2004), the molecular surface of SLL-2 is comparatively negatively charged, and

some limited parts in this cleft are positively charged (Figure 1C). Most of the sugar binding sites in SLL-2; an *N*-glycosylation site with an *N*-linked sugar, a carbohydrate binding site (site-1), and an additional sugar binding site (site-2) are assembled in this cleft (Figures 1A-C).

Two trimers of SLL-2 form a pseudo-intertwined, dumbbell-shaped hexameric molecule, with dimensions of 105 Å in length by 45 Å in diameter. Two globular domains are connected by three pairs of β-strands like hydrogen bonding interactions using extended N-terminal amino acid residues 1-5 (βN, Figures 1A and 1B). Two trimers are related by non-crystallographic two-fold axes in the SLL-2 hexameric structure. The average rmsd value of Cα atoms of the six monomers is 0.2 Å.

The crystals of the SLL-2-GalNAc complex (SLL-2G) diffract to 1.6 Å resolution, which is higher than that of SLL-2 crystal. The asymmetric unit of the SLL-2G crystal contains one monomer, and crystallographic symmetry-related molecules form the hexameric assembly (Figure 1C). The overall structures of the SLL-2G monomer and trimer are very similar to those of SLL-2; the average rmsd value for the Cα atoms of each monomer molecule and each trimer molecule between SLL-2 and SLL-2G are

both 0.3Å, respectively, whereas, that between hexamers is calculated as 1.1 Å. When one of the two trimers from SLL-2 is superimposed to the SLL-2G trimer, the other trimer of SLL-2 shows the movement of about 1.3 Å for the center of mass position with a rotation angle of about 7˚ counterclockwise against the other trimer of SLL-2G (Figure 1D). A slight shortening of the overall molecule length accompanied by this twist motion was observed in SLL-2G.

SLL-2 belongs to the H-type lectin family, which includes HPA, discoidins I and II from the slime mold *Dictyostelium discoideum* (Sanchez et al. 2006). H-type lectin is a hexameric or a trimeric molecule where an immunoglobulin-like β-sandwich fold consists of six β-strands exists in the monomer (HPA and discoidin I in Figures 1E and 1F) (Sanchez et al. 2006; Aragao et al. 2008; Mathieu et al. 2010). HPA is a homo hexamer consisting of two sets of trimers interconnected with three disulfide bonds between two monomers facing each other (Figure 1E) (Sanchez et al. 2006). On the other hand, discoidins I and II are intertwined trimeric molecules (Mathieu et al. 2010; Aragao et al. 2008), in which the monomer consists of the H-type lectin and the N-terminal domain connected by linker residues (discoidin I, Figure 1F). Since SLL-2

is a hexameric molecule with a pseudo-intertwining arrangement, its assembly has characteristics of both HPA and discoidins I and II.

Isoproteins in SLL-2 crystal

We previously reported the presence of three major isolectins in a sample of SLL-2, comprised of the 55th histidine, asparagine, and arginine in SLL-2a, 2b, and 2c (Supplementary Figure S1A, Jimbo et al. 2005). The electron density at the 55th amino acid for SLL-2G indicated the presence of His55, Asn55, and the possibility of Arg55 (Supplementary Figure S1B). This fact shows that at least two isolectins (SLL-2a and 2b) exist, and also suggests the possibility of the presence of SLL-2c in the SLL-2 crystal. The conformations of the main chains of the isoproteins, however, are quite similar that they can construct SLL-2G hexameric molecule with the space group of *I*432. As SLL-2b shares common amino acid residues largely with three other isoproteins, we have adopted the amino acid sequence of SLL-2b as the sequence of SLL-2 for crystal structure refinement.

N-glycosylation site of SLL-2 protein

The *N*-glycosylation site is observed clearly at Asn60 of each SLL-2 molecule. The first and second sugars (*N*-acetylglucosamine (GlcNAc) - GlcNAc) in the oligosaccharides are fitted reasonably into the electron density map of SLL-2 (Supplementary Figure S2). The amide nitrogen of the first GlcNAc associates with high electron density species tentatively fitted as a chloride, and the carbonyl oxygen forms a hydrogen bond with the amide group of the residue Asn76. The interactions of the amide-anion are not only seen in some protein coordinates registered in PDB, but also indicated by recent MD simulations (Algaer & Vegt, 2011) and *ab initio* calculations (Yu et al. 2010). Based on the MALDI-TOF MS results, one SLL-2 monomer carries *N*-linked glycans with a mass of approximately 2,500 Da, corresponding to more than 10 saccharides (Jimbo et al. 2005). Therefore, flexible oligosaccharides consisting of more than eight sugar units possibly exist at the unoccupied space in the SLL-2 crystal.

The sugar binding site in SLL-2G, site-1

In SLL-2G, the electron density for GalNAc was observed clearly in the deep pocket at the interface between the monomers, which we define as site-1 (Figures 1A and 1C). Site-1 is the position corresponding to the carbohydrate binding site in other H-type lectins (Figures 2A-C). Site-1 consists of the side chains of Asn56, Arg58 in the β4 strand, Trp78 and the 79th amino acid residue (Tyr or Ala) in the loop between β5 and β6 strands, Trp18' in the loop between β1' and β2' strands, Asp50' in the β3' strand, and Trp84' and Asn85' in the β6' strand.

The interactions between GalNAc and the SLL-2 protein in site-1 are; between the O3 atom of the galactopyranoside and the Nε1 atom of Trp78; the O4 atom and the Nε atom of Arg58, the Nε1 atom of Trp78, and the Oδ1 atom of Asp50'; the O6 atom and the Nδ2 atom of Asn56, the Nη2 and Nε atoms of Arg58, and the Oδ2 atom of Asp50'; the pyrane oxygen and the Nη2 atom of Arg58 (Figure 2A). In these residues, Arg58, Trp78, and Asp50' are completely conserved in the H-type lectin family (Figures

2A-2C and Supplementary Figure S1A). The side chain of Trp18', which is a unique residue among the H-type lectins, forms saccharide-aromatic residue interaction against the galactopyranoside ring. This observation is consistent with the previous report that the interaction between galactopyranoside and the aromatic residue using the relatively non-polar B-face of galactose, in which the C3-H, C4-H, C5-H, and C6-H atoms are embedded, is favored (Sujatha et al. 2004). Moreover, the electron density derived from the residue Tyr79, which is not conserved in three SLL-2 isomers, can interact with C1-H and C2-H atoms by van der Waals forces.

Site-1 in SLL-2

The electron density for site-1s in SLL-2 are not so clear as opposed to the result from the SLL-2G. They are classified in two groups, saccharides binding sites and others. Two of the six site-1s in SLL-2, however, can be fitted each with a galactopyranoside derivative, a GalNAc-containing oligosaccharide and a galactose (Figures 2D and 2E). Interestingly, residual electron density elongated in the direction of the putative

N-linked glycans is observed at site-1 with a GalNAc (Figure 2D). The interactions between galactopyranoside derivatives and the SLL-2 protein in these site-1s are almost identical with those of SLL-2G. The nature of this elongated moiety is not clear, but either external sugars (from the purification process, or even coral origin) or internal sugars (*N*-linked glycans of neighboring lectin molecules or even from the lectin itself) are conceivable.

Three of six site-1s in SLL-2 accommodate an MPD molecule (Figure 2F). The framework of the MPD molecule can be superimposed onto the galactopyranoside ring of the other site-1s, that is; two hydroxyl groups of the MPD molecule (O2 and O4) are located at the positions corresponding to the O4 atom of the galactopyranoside ring and the O6 atom of the hydroxymethyl moiety. Notably, the indole ring of Trp18', which makes interaction with the galactopyranoside ring when a galactose derivative is bound to site-1, flipped toward the outside of the cleft (Figure 2F).

The remaining one of the six site-1s in SLL-2 possesses electron density assignable to a water oxygen atom, which is located at the position corresponding to the O6 atoms of galactopyranoside derivatives in other site-1s (Figure 2G). The direction of the side

chain of Trp18' was oriented toward the outside of the cleft. The discontinuous electron densities around the Cβ atom of the 79th residue are observed (Figure 2G). Due to the ambiguous electron densities, the conformation of this residue is not defined. These observations suggest to us that site-1s in SLL-2 can accommodate structurally diverse molecules, and that SLL-2 should be able to maintain the hexameric molecule stably with or without the sugar ligands in site-1s.

Saccharides recognition by SLL-2 assemblage

SLL-2 selectively binds to galacto- over gluco-saccharides with α-anomers. The complex is formed with the saccharides of galacto-configuration sugar at the nonreducing end in site-1, that is located between the side chains of Trp18' and Tyr/Ala79 (Figure 2A and Supplementary Figures S3A and S3B). Here we showed atomic interactions between SLL-2 sugar binding residues and galactosaccharide. The binding between galactosyl sugar and SLL-2 is stabilized by interactions between the axial O4 atom of the galactosyl moiety and the side chains of Asp50', Arg58, and Trp78.

These interactions are reduced when GlcNAc is substituted for GalNAc; that is, the O4' atom of the GlcNAc loses three interactions to create two new interactions with the side chain of Asp50' and Asn85' (Figure 2A and Supplementary Figure S3A). These recognition systems for galactopyranoside are also observed in other H-type lectins. In addition, Trp18' in SLL-2, which is the unique residue among H-type lectins, contributes to the saccharide-aromatic residue interactions. Because the O4 atom in the GlcNAc would reduce the C4-H interaction between saccharide and Trp18', Trp18' would contribute to the recognition of galactosyl sugar. Besides, our previous FAC analyses indicated that the association constant of α-GalNAc for SLL-2 is more than three times larger than that for HPA (Jimbo et al. 2013), even though interactions between two residues (Gly24, Asp26) and α-GalNAc observed in HPA are missing in SLL-2G. Trp18' might contribute not only to the selectivity towards the galactopyranoside, but also to the specificity difference of α-GalNAc between SLL-2 and HPA by constructing the environment of site-1 as described below.

The atomic structure also shows that SLL-2 prefers α- over β- anomer, as suggested by its binding preference for the Forssman antigen. In the crystal structures, van der

Waals contact between the equatorial position of C1 and Tyr79 was found, where O1' atom of the β-anomer would occupy the space (Supplementary Figure S3A). The anomeric restriction by the residue corresponding to Tyr79 in SLL-2 is also observed at the corresponding position in the HPA protein; the steric hindrance of His84 appeared to inhibit the binding of β-anomer (Supplementary Figure S3C, Sanchez et al. 2006). Discoidins I and II, which can accommodate both anomers, have no aromatic residue at the corresponding position. Notably, the saccharide was observed in all but one of the SLL-2 site-1s in which the electron density of Tyr79 is disordered. We therefore conclude that van der Waals contacts between the aromatic residue and C1-H-C2-H in the galactopyranoside might contribute to selectivity toward the α-anomer in H-type lectin.

The additional sugar binding site, site-2

The additional sugar binding sites, site-2s, are observed at the top and the bottom of the hexameric molecule (Figures 1A and 1C). Site-2 is comprised of the residues Asp52',

the 53'rd amino acid (Asn or Ser), the 54'th amino acid (Ser or Thr), and the residue Asn56 (Figure 3A and Supplementary Figure S3B). Though the electron density map of site-2 is clearly observed, and a pyranose ring was assignable on the basis of the size and the shape of the electron density; however, it was difficult to define the substitution groups and orientation of the ring, due presumably to the amino acid sequence divergence at the 54'th (Ser or Thr) residue which might affect the binding. The residues Asp52', the 53'rd amino acid (Asn or Ser), the 54'th amino acid (in the case of Thr of SLL-2c), and Asn56, and an MPD molecule used as crystallization reagent, are within the hydrogen bond distance from site-2 (Figure 3A). The MPD molecule is located along the line between site-1 and site-2, and bridges these sites. In addition, near site-2, a chloride ion which can interact with the side chain of Arg58 is located. Consequently, both site-1 and site-2 are involved in the network.

In the crystal structure of SLL-2, the electron densities for site-2s are observed at four of the six equivalent positions. Two other sites, however, are unoccupied since the atoms of neighboring molecules disrupt the binding of the ligand due to steric hindrance induced by the crystal packing. This observation suggests that the interaction between

SLL-2 and the saccharide at site-2 is of low-affinity. The fact that electron densities for site-2s are observed even in SLL-2, which contains no external saccharides under the crystallization condition, and the fact that residual electron densities are observed when a pyranose molecule was fitted in SLL-2G, together suggest that an external sugars (from purification step or from coral), internal sugars (*N*-linked sugars of SLL-2), or GalNAc added in the crystallization procedure (in the case of SLL-2G) is bound to site-2.

Comparison of specificity loops with H-type lectins

In H-type lectins, the first and third loops (the loop between β1 and β2, and that between β3 and β4) are supposed to be the loops that control the specificity of sugar binding site (Mathieu et al. 2010). In SLL-2, these loops contribute to form site-1 and site-2 (Supplementary Figures S3B-D).

The first loop in H-type lectins is the most deviated region that contains various amino acid residues of unequal length (Supplementary Figure S1A). In SLL-2, the

N-terminal part of the first loop, in which Trp18' residue is contained, spreads to site-1 of the neighbor molecule (Figure 2A and Supplementary Figure S3B). The side chains of Trp18', Tyr79 (in the case of SLL-2a and 2b) in β5 - β6 loop, and Trp84' in β6 associate to make site-1 deeper and narrower. In HPA, the first loop extends in the direction perpendicular to the immunoglobulin-like strands, and two residues in this loop, Gly24 and Asp26, interact with *N*-acetyl group and O3 of GalNAc in site-1 (Figure 2B and Supplementary Figure S3C). Site-1 in HPA is restricted by the first loop and the side chains of His84 and Tyr89', which are corresponding to Tyr79 and Trp84' residues in SLL-2. Due to the absence of aromatic residue corresponding to Trp18' in SLL-2, HPA contains site-1 that is slightly wider than that of SLL-2 (Supplementary Figure S3C). In discoidins, the first loop is comparatively short and does not play a role to interact with sugars in site-1 (discoidin I in Supplementary Figure S3D). Tyr241', which is located at a near position corresponding to Trp18' in SLL-2, shows a little difference in relative orientation to GalNAc (Figure 2C and Supplementary Figure S3D). Accordingly, the short first loop and the aromatic residue in different position and orientation in discoidin I contribute to form the

wide-spread site-1 in H-type lectins. The absence of the aromatic residue corresponding to Trp84' in SLL-2 allows to extend a cleft in the direction of C5 of GalNAc (Supplementary Figure S3D).

Site-2 in SLL-2 consists of amino acids involved in the third loop (Supplementary Figure S3B). The main chain of the third loop in SLL-2 is well superposed to that in HPA, and the properties of amino acid residues in the third loops are similar between SLL-2 and HPA. On the other hand, the molecular surfaces around site-2 position in both discoidins are protruded due to the differences of the length of chains and their conformation (discoidin I in Supplementary Figure S3D). If site-2 is the subsite for interaction with the extension of oligosaccharide chains in site-1 as described below, the third loop might affect the affinity for oligosaccharide chains accommodated in H-type lectins.

The possible Forssman antigen binding mode in SLL-2 protein

The frontal affinity chromatography experiments for SLL-2 demonstrated that the

Forssman antigen is efficiently bound to SLL-2 (Jimbo et al. 2013). The HPA molecule also exhibited Forssman antigen binding affinity. The HPA crystal structure complexed with the antigen disaccharide (PDB code: 2cgy) (Lescar et al. 2007) indicated that the disaccharide extends its reducing end against the trimer surface. The superposition of the HPA-Forssman antigen disaccharide model over SLL-2G reveals that the GalNAc in SLL-2G and nonreducing GalNAc in the HPA-Forssman antigen disaccharide are located in the same position with the same trajectory, allowing the second GalNAc residue in the HPA-Forssman antigen disaccharide extend toward the position of the MPD molecule observed in SLL-2G (Figure 3B). The conformations of residual three sugars of the Forssman antigen pentasaccharide in its complex with SLL-2 are unknown. However, we have recently demonstrated that the binding affinity of SLL-2 was increased as the length of the Forssman antigen extended from the tri- to the pentasaccharide (Tanaka et al. 2013). This fact suggests that increasing chain length from tri to pentasaccharide tends to prompt interaction between SLL-2 and the internal sugars of pentasaccharide. The present results together suggested a possible binding interaction between SLL-2 and the Forssman antigen pentasaccharide

as follows; 1) the nonreducing GalNAc residue of the Forssman antigen binds to site-1 in a manner that GalNAc binds to SLL-2G; 2) the pentasaccharide extends out from the core of a protein in a manner found in MPD molecule in SLL-2G; 3) the pentasaccharide chain bends at a corner of the SLL-2 molecule because of the α-(1,4)-glycosidic bond between the third and fourth sugars; and 4) the reducing end of the pentasaccharide binds to site-2 (Figure 3C). The distance between site-1 and site-2, approximately 10 - 11 Å, corresponded well with the distance between the nonreducing terminal GalNAc residue and the reducing end of pentasaccharide. The crystal structure of the complex between SLL-2 and Forssman antigen-like polysaccharide will enable us to discuss their detailed binding manner.

Materials and Methods

Crystallization and data collection

SLL-2 was purified from the coral *Sinularia lochmodes* according to the protocol

reported by Jimbo (Jimbo et al. 2005). The purified SLL-2 was prepared at a concentration of 20 mg/mL in 50 mM Tris buffer, pH 8.0. The rhombic plate crystals were obtained from 2-methyl-2,4-pentanediol (MPD) solutions containing various additive reagents. The best crystals of SLL-2 were grown at 293 K by the sitting drop vapor diffusion method from a droplet consisting of equal volumes of protein solution and reservoir, which consists of 40% MPD and 0.2 M calcium formate in 0.1 M Bis-Tris buffer, pH 6.5. The space group of SLL-2 crystal was $P2_1$ with cell constants of a = 52.0 Å, b = 189.9 Å, c = 57.1 Å, and β = 110.6˚. The screenings of the crystallization condition to obtain SLL-2-sugar complex crystals were also carried out in the presence of five-fold molar excess of D-galactose or GalNAc, respectively. No crystals were obtained from the solutions of SLL-2 and D-galactose. The co-crystals of SLL-2 and GalNAc (SLL-2G) were grown using a solution containing 45% MPD, 0.2 M calcium chloride, and 0.1M Bis-Tris buffer, pH 5.5. SLL-2G crystals were cubic shaped and belonged to space group $I432$, with unit-cell parameters of $a = b = c$ = 154.2 Å. All of the obtained crystals were soaked in the precipitant solutions, and then submitted to flash-freezing directly within a cold nitrogen-gas stream (-173˚C).

Diffraction data were collected using ADSC Quantum 270 at the BL-17A beam line and ADSC Quantum 210 at NW12 beam line, Photon Factory (KEK; Tsukuba, Japan). The intensity data were processed by using the HKL2000 program package (Otwinowski & Minor, 1997), and truncated to structure factors using the program CCP4 (Collaborative Computational Project, Number 4 1994). The data statistics are shown in Supplementary data, Table I.

Structure determination and refinement

Molecular replacement was performed for phase determination with the program MOLREP (Vagin & Teplyakov 1997) in the CCP4 package. The atomic coordinates of the trimeric part of hexameric HPA protein (PDB code: 2ccv) were used as a search model (Sanchez et al. 2006). The structural model was built by manual fitting to the electron density map by using the program TURBO FRODO (Roussel & Cambillau 1992). All structural refinements were carried out using the program CNS (Brunger et al. 1998). The stereochemical quality of the final models was assessed using the

program PROCHECK (Laskowski et al. 1993). The molecular models in the figures were drawn by using the program PyMOL Molecular Graphic System (Version 1.2r3pre, Schrödinger, LLC). Data statistics are also summarized in Supplementary data, Table I.

PDB references

The atomic coordinates and structure factors of SLL-2 and SLL-2G have been deposited in the Protein Data Bank, with the accession code of 3WMP and 3WMQ.

Acknowledgements

We would like to thank the staff members at the beamlines of Photon Factory for their help with X-ray data collections with synchrotron radiation.

Conflict of Interest

None declared.

Abbreviations

SLL-2, lectin from *Sinularia lochmodes*; HPA, *Helix pomatia* agglutinin; GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; MPD, 2-methyl-2, 4-pentanediol; Forssman antigen (GalNAcα1-3GalNAcβ1-3Galα1-4Galβ1-4Glc-ceramide); DBA, *Dolichos biflorus* agglutinin; SLL-2G, SLL-2 complex with GalNAc;

References

Adams LM, Cumbo VR, Takabayashi M. 2009. Exposure to sediment enhances primary acquisition of Symbiodinium by asymbiotic coral larvae. *Mar. Ecol. Prog. Ser.* 377:149-156.

Algaer EA, van der Vegt, NFA. 2011. Hofmeister ion interactions with model amide compounds. *J. Phys. Chem.* B115:13781-13787.

Aragao KS, Satre M, Imberty A, Varrot A. 2008. Structure determination of discoidin II from *Dictyostelium discoideum* and carbohydrate binding properties of the lectin domain. *Proteins* 73:43-52.

Baker AC, Starger CJ, McClanahan TR, Glynn PW. 2004. Coral reefs: Corals' adaptive response to climate change. *Nature* 430:741.

Baker AC, Glynn PW, Riegl B. 2008. Climate change and coral reef bleaching: An ecological assessment of long-term impacts, recovery trends and future outlook. *Est. Coast. Shelf Sci.* 80:435-471.

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* 98:10037-10041.

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, *et al*. 1998. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Cryst.* D54:905-921.

CCP4 (Collaborative Computational Project, Number 4). 1994. The CCP4 suite: programs for protein crystallography. *Acta Cryst.* D50:760–763.

Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. 2004. PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 32:W665-W667.

Fukuzawa M, Ochiai H. 1996. Molecular cloning and characterization of the cDNA for discoidin II of *Dictyostelium discoideum*. *Plant Cell Physiol*. 37:505-514.

Jimbo M, Yanohara T, Koike K, Sakai R, Muramoto K, Kamiya H. 2000. The D-galactose-binding lectin of the octocoral *Sinularia lochmodes*: characterization

and possible relationship to the symbiotic dinoflagellates. *Comp. Biochem. Physiol. B Biochem. Mol. Biol*. 125:227-236.

Jimbo M, Koike K, Sakai R, Muramoto K, Kamiya H. 2005. Cloning and characterization of a lectin from the octocoral *Sinularia lochmodes*. *Biochem. Biophys. Res. Commun*. 330:157-162.

Jimbo M, Suda Y, Koike K, Nakamura-Tsuruta S, Kominami J, Kamei M, Hirabayashi J, Sakai R, Kamiya H. 2013. Possible involvement of glycolipids in lectin-mediated cellular transformation of symbiotic microalgae in corals. *J. Exp. Mar. Biol. Ecol.* 439:129–135.

Koike K, Jimbo M, Sakai R, Kaeriyama M, Muramoto K, Ogata T, Maruyama T, Kamiya H. 2004. Octocoral chemical signaling selects and controls dinoflagellate symbionts. *Biol. Bull.* 207:80-86.

Laskowski RA, MacArthur MA, Moss DS, Thornton JM. 1993. PROCHEK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst*. 26:283-291.

Lescar J, Sanchez JF, Audfray A, Coll JL, Breton C, Mitchell EP, Imberty A. 2007. Structural basis for recognition of breast and colon cancer epitopes Tn antigen and Forssman disaccharide by *Helix pomatia* lectin. *Glycobiology* 17:1077-1083.

Lewis CL, Coffroth MA. 2004. The acquisition of exogenous algal symbionts by an octocoral after bleaching. *Science* 304:1490-1492.

Mathieu SV, Aragao KS, Imberty A. Varrot A. 2010. Discoidin I from *Dictyostelium doscoideum* and interactions with oligosaccharides: specificity, affinity, crystal structures, and comparison with discoidin II. *J. Mol. Biol*. 400:540-554.

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

Poole S, Firtel RA, Lamar E, Rowekamp W. 1981. Sequence and expression of the discoidin I gene family in *Dictyostelium discoideum*. *J. Mol. Biol.* 153:273-289.

Roussel A, Cambillau C. 1992. TURBO FRODO program, Mountain View CA: Silicon Graphics Geometry Partners Directory

Sampayo EM, Ridgway T, Bongaerts P, Hoegh-Guldberg, O. 2008. Bleaching susceptibility and mortality of corals are determined by fine-scale differences in symbiont type. *Proc. Natl. Acad. Sci. USA* 105:10444-10449.

Sanchez JF, Lescar J, Chazalet V, Audfray A, Gagnon J, Alvarez R, Breton C, Imberty A, Mitchell EP. 2006. Biochemical and structural analysis of *Helix Pomatia* agglutinin (HPA): A hexameric lectin with a novel fold. *J. Biol. Chem*. 281:20171-20180.

Sujatha MS, Sasidhar TU, Balaji PV. 2004. Energetics of galactose- and glucose-aromatic amino acid interaction: Implications for binding in galactose-specific proteins. *Protein Sci.* 13:2502-2514.

Tanaka H, Takeuchi R, Jimbo M, Kuniya N, Takahashi T. 2013. Synthesis and biological evaluation of the Forssman antigen pentasaccharide and derivatives by a one-pot glycosylation procedure. *Chem. Eur. J*. 19:3177-3187.

Vagin AA, Teplyakov A. 1997. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst*. 30:1022-1025.

Yu H, Mazzanti CL, Whitfield TW, Koeppe RE, Andersen OS, Roux B. 2010. A combined experimental and theoretical study of ion solvation in liquid N-Methylacetamide. *J. Am. Chem. Soc.* 132:10847-10856.

FIGURE LEGENDS

Figure 1. The overall structure of SLL-2. (A) The monomer conformation of SLL-2. The Fo-Fc electron density maps around site-2 contoured at 3 σ are drawn. (B) The overall structures of the hexameric SLL-2. The top (upper) and front (lower) views were drawn. Two braces indicate each unit of trimeric globular form of SLL-2, respectively. The molecules in three site-1s belonging to the trimeric globular form 1 (two MPDs and one GalNAc moiety) are labeled by black letters, and the molecules in the remaining site-1s belonging to the trimeric globular form 2 (one MPD and one galactose) are labeled by purple letters. *N*-linked sugars and MPDs, the GalNAc moiety, and the galactose in site-1s are colored in white, green, yellow, and orange sticks. The electron density maps for site-2s are omitted for clarity. (C) The

molecular electrostatic potential distribution of the SLL-2G (the top and front views) calculated with SLL-2b hexamer molecule. Positive and negative potential areas colored by potential on solvent accessible surfaces are shown in blue and red, respectively (Baker et al. 2001; Dolinsky et al. 2004). *N*-linked sugars and sugars in site-1s are drawn as white and yellow stick models. The electron density maps for site-2s contoured at 3 σ are indicated by black mesh. (D) Superposition of one trimeric globular domains (the lower part of the lower figure, drawn in ribbon models) of SLL-2 and that of SLL-2G. The other trimeric domains of SLL-2 and SLL-2G were indicated by green and red lines, respectively. The upper figure is the top view, and the lower one is the front view. (E) The ribbon model of HPA complexed with GalNAc (PDB code: 2ccv) (Sanchez et al. 2006). (F) The ribbon model of discoidin I with GalNAc (PDB code: 2w95) (Mathieu et al. 2010).

Figure 2. Site-1s in SLL-2 and H-type lectins. Electron density maps in SLL-2G and SLL-2 are sugar-omitted Fo-Fc maps contoured at 3 σ. The dashed lines indicate interactions. (A) Site-1 in SLL-2G with GalNAc. (B) Site-1 in HPA with GalNAc.

(C) Site-1 in discoidin I with GalNAc. (D) Site-1 in SLL-2 with GalNAc moiety. (E) Site-1 in SLL-2 with galactose. (F) Site-1 in SLL-2 with MPD. (G) Site-1 in SLL-2 with water molecule. Site-1 and the 79th residue omitted Fo-Fc electron density map is shown.

Figure 3. Binding model of oligosaccharides to SLL-2. (A) The environment of site-2 in SLL-2G. The Fo-Fc electron density map at site-2 contoured at 4 σ is shown with tentatively fitted pyranose ring (orange stick). The black dashed lines indicate interactions. (B) Superposition of the Forssman antigen disaccharide of HPA (PDB code: 2cgy) on site-1 of SLL-2. GalNAc in site-1 and MPD molecule in SLL-2G are depicted as yellow and green stick models with the cartoon model of SLL-2G. Forssman antigen disaccharide of HPA is drawn in a white stick model. (C) One of the possible binding modes of Forssman antigen pentasaccharide (white stick model) to SLL-2. GalNAc, MPD, and electron density for site-2 are also drawn in yellow and green sticks, and gray mesh.

Legends for Supplementary Figures

Supplementary Figure S1.

Sequences of H-type lectins.

(A) Structure-based sequence alignment of SLL-2 isomers and H-type lectins. The common residues are colored red and cyan. The secondary structure elements of SLL-2 are displayed above the sequence. The lines colored in magenta and orange indicate the first and the third loops. (B) The electron density map of SLL-2G at the 55th amino acid residue. The blue and orange meshes indicate the 2Fo-Fc map contoured at 2 σ and the Fo-Fc map contoured at 3 σ of SLL-2b. Cyan, green, and magenta stick models show residues of SLL-2a, 2b, and 2c, respectively.

Supplementary Figure S2.

N-glycosylation site in SLL-2G.

N-linked sugar is represented with its omitted Fo-Fc electron density map contoured at 3 σ. A chloride ion is represented by the gray ball.

Supplementary Figure S3.

Environments of site-1s.

(A) The saccharide recognition in site-1 in SLL-2. Two covalent bonds between epimeric carbons and oxygen atoms (O1' and O4') that belonging to the other epimers are colored cyan. The interactions between SLL-2 and the O4 atom in α-GalNAc are drawn by orange lines, and those with the O4' atom in GlcNAc are drawn by blue lines.

(B, C, D) Protein surfaces around site-1s and specific loops of SLL-2G, HPA, and discoidin I. The first and the third loops are colored in magenta and orange. The aromatic residues noted in the text are drawn in stick model. *N*-linked sugar is drawn in white sticks. The Fo-Fc electron density map contoured at 4 σ shown for site-2 in SLL-2G is also drawn as mesh to clarify its location.

Figure 1

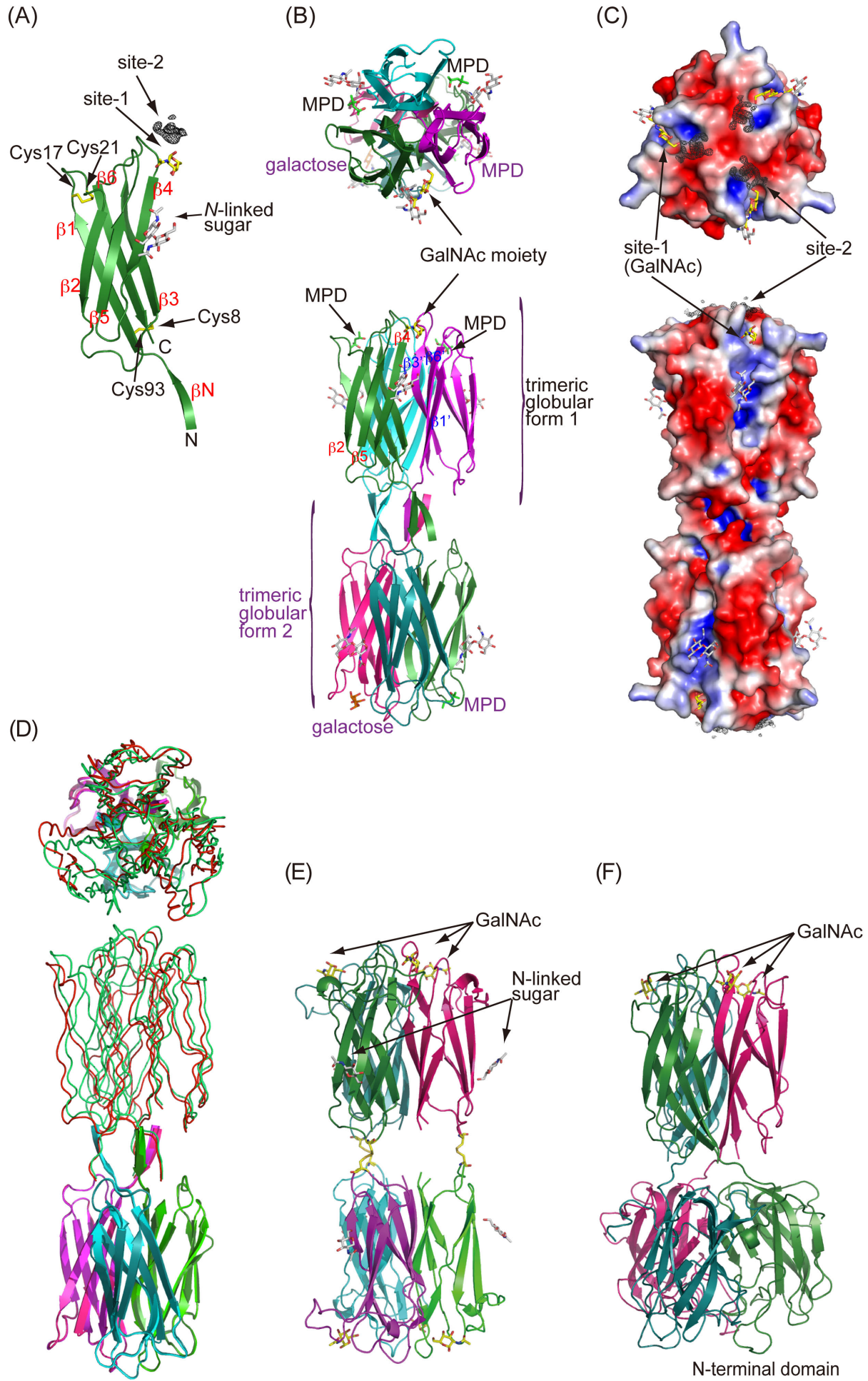

(A)
site-2
site-1
Cys17
Cys21
β6
β4
N-linked sugar
β1
β2
β3
β5
Cys8
C
Cys93
βN
N
(B)
MPD
MPD
galactose
MPD
GalNAc moiety
MPD
MPD
β4
β3'N
β1'
β2
β5
trimeric globular form 1
trimeric globular form 2
MPD
galactose
(C)
site-1 (GalNAc)
site-2
(D)
(E)
GalNAc
N-linked sugar
(F)
GalNAc
N-terminal domain

Figure 2

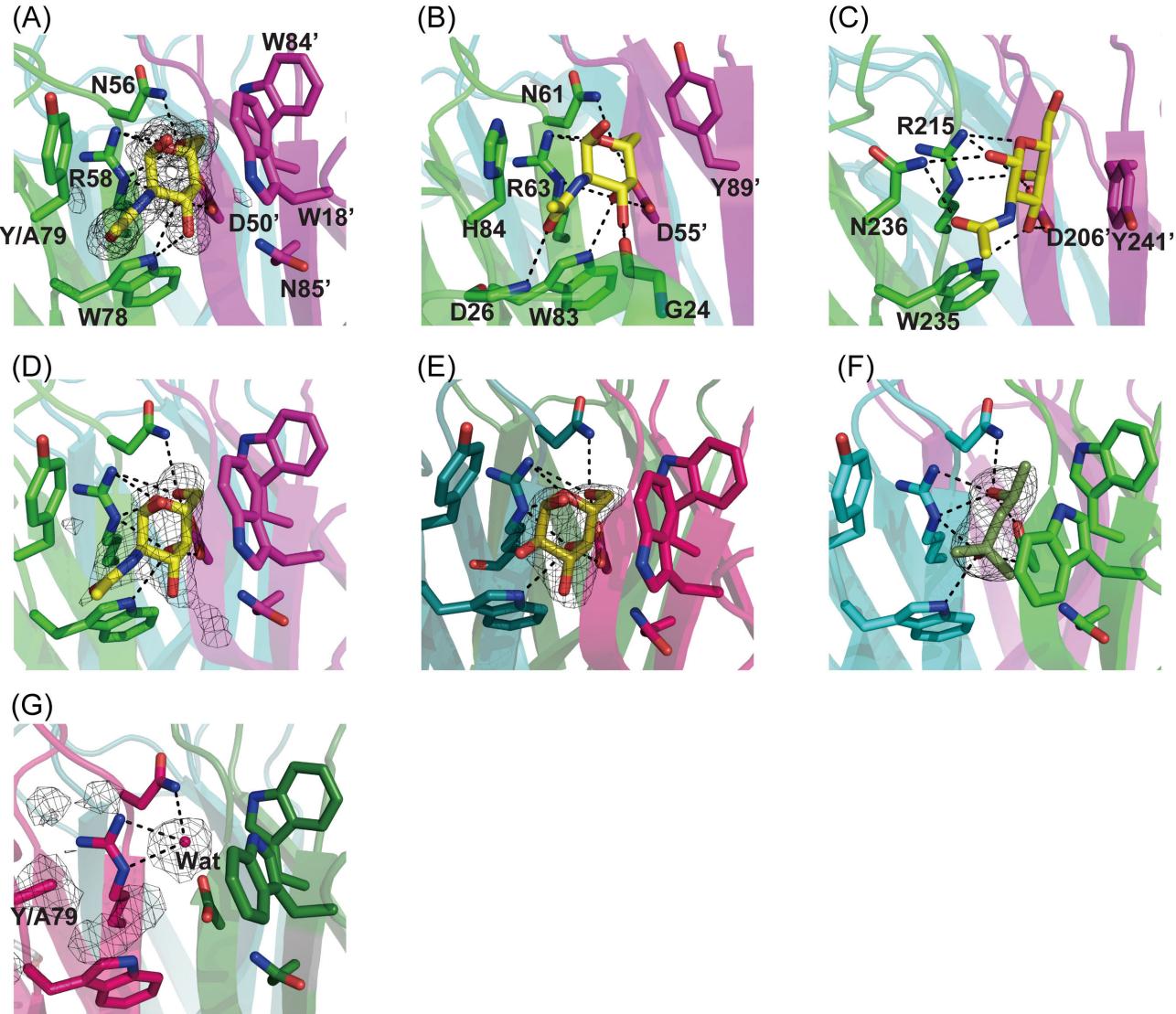
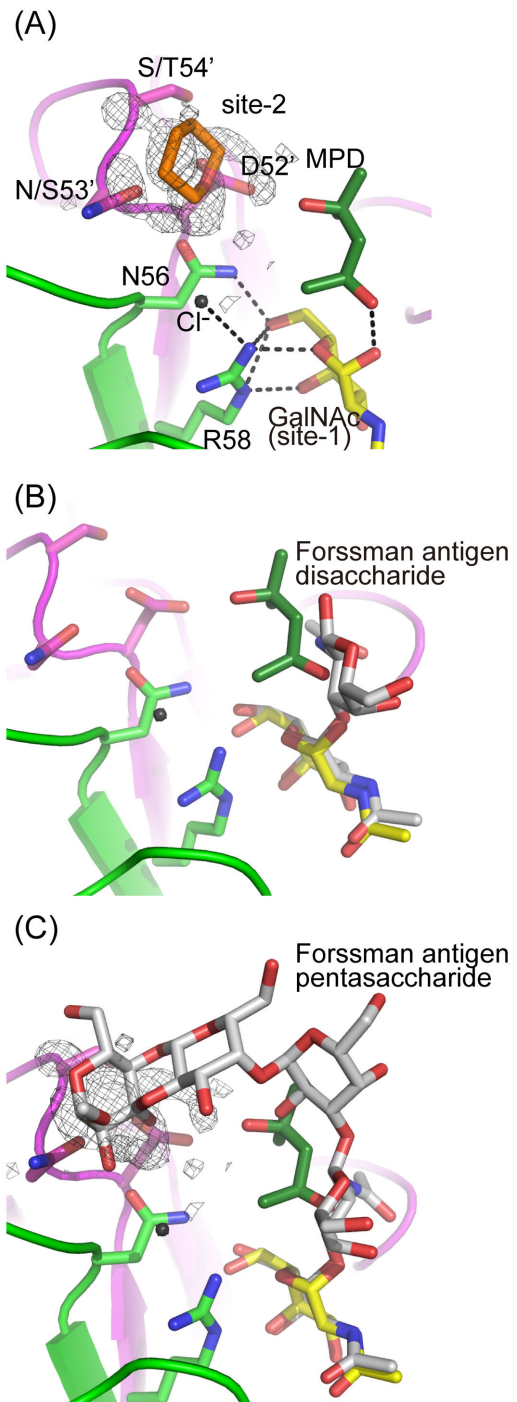
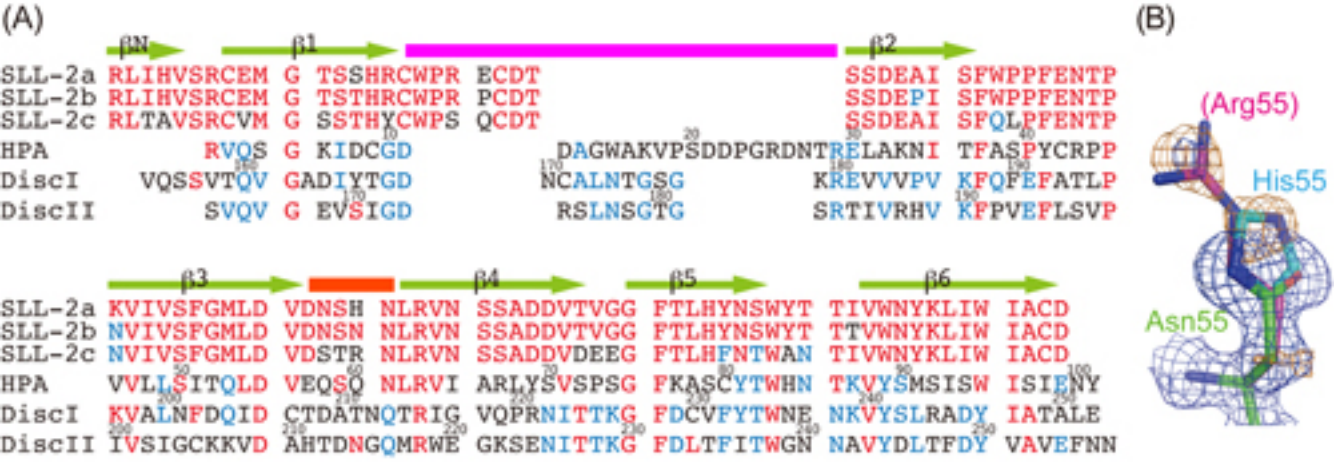

(A)
W84'
N56
R58
Y/A79
D50'
W18'
N85'
W78
(B)
N61
R63
H84
Y89'
D55'
D26
W83
G24
(C)
R215
N236
D206'
Y241'
W235
(D)
(E)
(F)
(G)
Wat
Y/A79

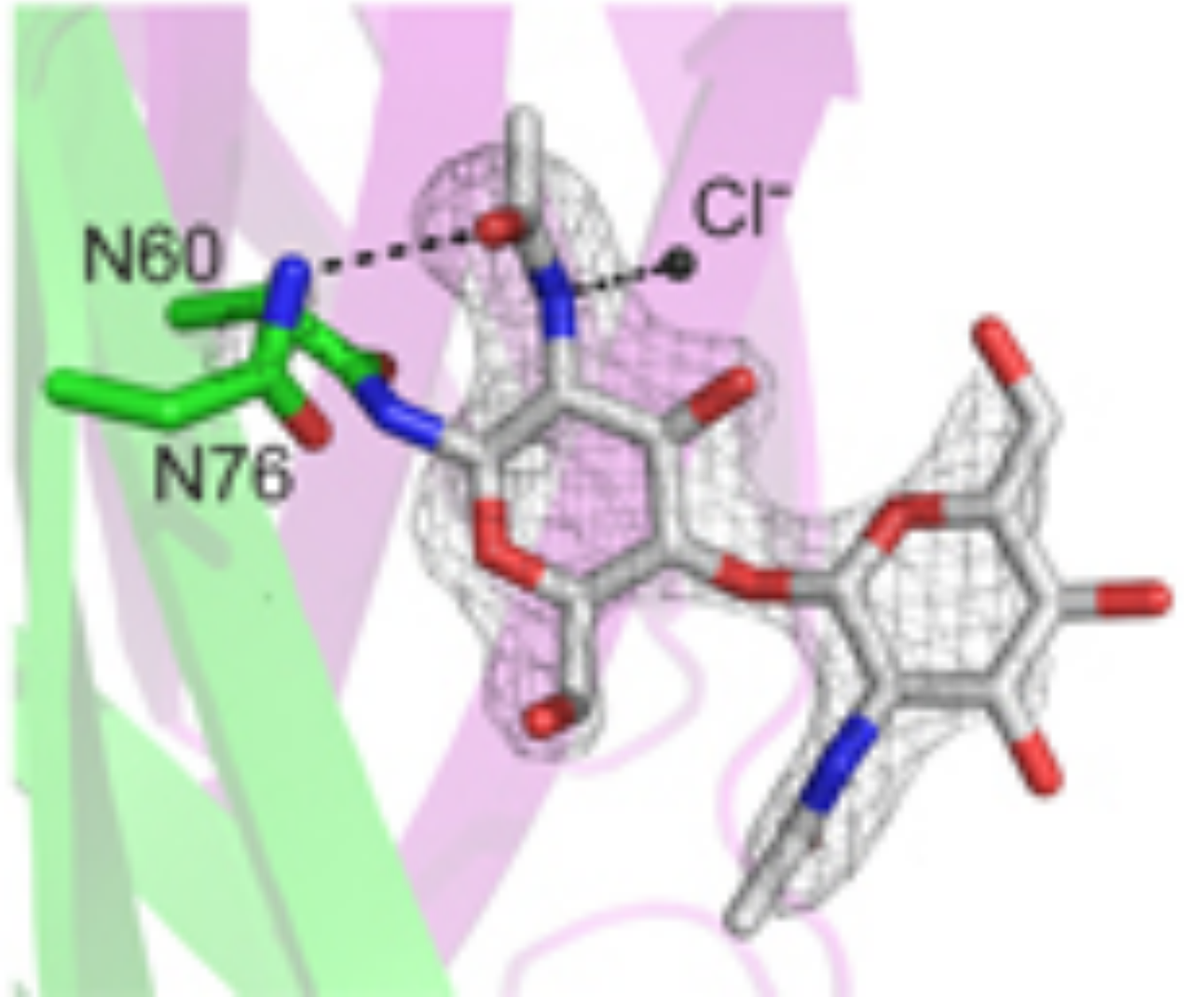
Figure 3


(A)
S/T54'
site-2
D52'
MPD
N/S53'
N56
Cl-
R58
GalNAc
(site-1)
(B)
Forssman antigen
disaccharide
(C)
Forssman antigen
pentasaccharide

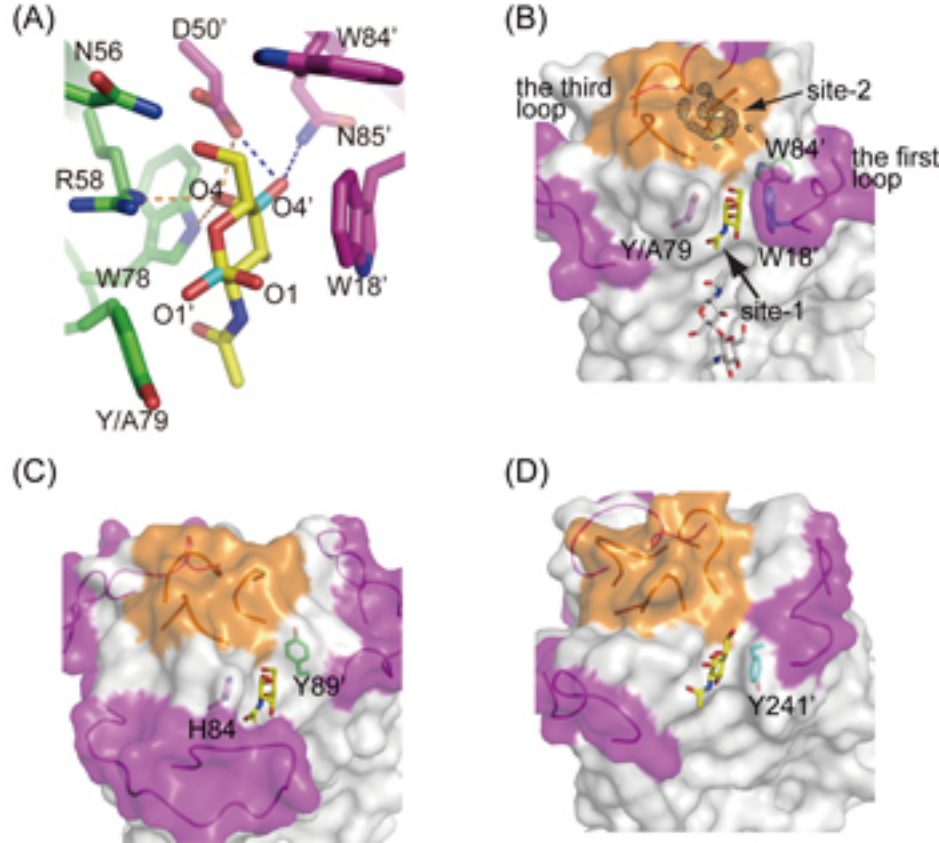
Supplementary Figure S1



Supplementary Figure S2



Supplementary Figure S3



Supplementary data

Table I. Data collection and refinement

Data collection statistics		
Protein	SLL-2	SLL-2G
PDB code	3wmp	3wmq
Wavelength	1.000	1.000
Space group	$P2_1$	$I432$
Cell dimensions (Å)		
a	52.0	154.2
b	189.9	
c	57.1	
β	110.6	
No. of molecules		
per asymmetric unit	6	1
Resolution (Å)	100.0 – 2.0	100.0 – 1.6
Measurements	338,957	490,810
Unique reflections	69,338	41,305
R_{merge}[a]	0.067 (0.308)[b]	0.079 (0.581)
Multiplicity	4.9 (4.8)	11.9 (11.5)
$I/\sigma(I)$	23.8 (3.1)	45.9 (5.0)
Completeness (%)	99.4 (98.8)	99.8 (100.0)
Overall B factor from		
Wilson plot (Å^2)	26.2	13.1
Refinement statistics		
Resolution range (Å)	100.0 – 2.0	100.0 – 1.6
No. of reflections	68,817	39,973
R_{work}/R_{free}[c,d] (%)	0.251/0.271	0.212/0.222
No. of atoms		
Protein/sugars/waters/others	4,529/195/121/42	759/43/78/21
Average *B* factors		
Protein/sugars/waters/others (Å^2)	45.4/60.4/47.6/52.7	17.2/20.2/29.6/36.1
R. m. s. d. from ideal geometry		
Bond length (Å)	0.008	0.007
Bond angles (°)	1.400	1.600
Ramachandran plot		
Most favored/allowed/outliers (%)		
	90.25/9.75/0.0	92.7/7.3/0.0

[a] $R_{merge} = \Sigma (|I-\langle I\rangle|)/\Sigma (I)$

[b] Numbers in parentheses refer to the highest-resolution shell, 2.03-2.00 Å for SLL-2, and 1.63-1.60Å for SLL-2 - GalNAc complex (SLL-2G).

[c] $R = \Sigma||F_{obs}|-|F_{calc}||/ \Sigma|F_{obs}|$

[d] R_{work} is calculated from a set of reflections in which 5% of the total reflections have been randomly omitted from the refinement and used to calculate R_{free}.