Introduction

Galectins, a family of sugar-binding proteins found in animals and fungi, are characterized by their evolutionarily conserved carbohydrate-recognition domains (CRDs) (Kasai and Hirabayashi 1996; Cooper 2002; Hirabayashi et al. 2002). Through their interactions with glycoconjugates that contain β-galactoside structures, such as Galβ1-4GlcNAc, galectins play important roles in various biological events such as development, immunity and cancer (Yang et al. 2008; Boschker et al. 2011). Galβ1-4GlcNAc binding activity, and the eight amino-acid residues of the CRD that are important for the recognition of this disaccharide, are conserved across the galectin family (Hirabayashi et al. 2002). Among vertebrate species, the major glyco-epitope for galectins is thought to be the Galβ1-4GlcNAc disaccharide unit, although galectins can also bind other disaccharide units containing β-galactoside structures (Iwaki et al. 2011). Conversely, among invertebrate species, the major glyco-epitope for galectins remains obscure: Multiple Galβ1-4GlcNAc units, which are the ligands of vertebrate galectins, have rarely been observed in species such as Caenorhabditis elegans and Drosophila melanogaster (Paschinger et al. 2008; ten Hagen et al. 2009).

Recently, however, the endogenous ligands for the major C. elegans galectin LEC-6, and LEC-10 were reported to be N-glycans containing Galβ1-4Fuc disaccharide units (Takeuchi et al. 2008, 2009; Maduzia et al. 2011): These N-glycans localize in the intestinal lumen, the grinder of the pharynx and the coelomocyte, where they function in the proper localization of the ligands of LEC-6 (Maduzia et al. 2011); They are also recognized by Coprinopsis cinerea galectin (CGL2) and this recognition resulted in the nematotoxicity of CGL2 (Butschi et al. 2010). In the genome of C. elegans, >10 galectin genes have been reported (Nemoto-Sasaki et al. 2008). At least three of them, LEC-1, LEC-6 and LEC-12a (DC2.3a), bind Galβ1-4Fuc in preference to Galβ1-4GlcNAc (Takeuchi et al. 2009, 2011; Nemoto-Sasaki et al. 2011). Furthermore, seven galectin proteins, LEC-1, -2, -4, -6, -9, -10 and -12a, were specifically isolated from C. elegans extracts using a Galβ1-4Fuc immobilized adsorbent (Takeuchi et al. 2011). Therefore, the major endogenous glyco-epitope for C. elegans galectins could be the Galβ1-4Fuc disaccharide unit, although the endogenous glycan ligand of LEC-8 does not contain such a disaccharide structure (Ideo et al. 2009).

Galβ1-4Fuc is a unique disaccharide unit, which has only been found in N-glycans of protostomia (Zhang et al. 1997; Takahashi et al. 2003, Wuhrer et al. 2004; Paschinger et al. 2008; Nemoto-Sasaki et al. 2008; ten Hagen et al. 2009).
A galactosyltransferase (GALT-1) responsible for synthesis of the Gal\(\beta\)-1-4Fuc disaccharide unit in *C. elegans*, was recently reported; Database searches revealed that GALT-1 homologs also exist in vertebrates (*Xenopus* and *Danio rerio*), invertebrates (*Caenorhabditis* and *Drosophila*), plants and protozoa (*Cryptosporidium*), but not in mammalian species (Titz et al. 2009). Although the presence of GALT-1 homologs does not prove the presence of the Gal\(\beta\)-1-4Fuc disaccharide unit, the existence of GALT-1 homologs suggests a role in recognition of glycan ligands not only by *C. elegans* galectins, but also by other galectins such as CGL2, which also has Gal\(\beta\)-1-4Fuc binding ability (Batschi et al. 2010). Therefore, it is possible that galectins in other species might also recognize this disaccharide unit, and the interaction between galectin and Gal\(\beta\)-1-4Fuc might be biologically important in species other than *C. elegans*.

In this study, we elucidated the molecular mechanism of the interaction between *C. elegans* LEC-6 and Gal\(\beta\)-1-4Fuc, which differs in important details from the interactions between vertebrate galectins and the Gal\(\beta\)-1-4GlcNAc unit (Takeuchi et al. 2009). We determined the crystal structure of LEC-6 in complex with galactose-\(\beta\)1,4-fucose (Gal\(\beta\)-1-4Fuc) at 1.5 Å resolution, and performed frontal affinity chromatography (FAC) analysis of Glu67 mutants to compare the affinities of wild-type and mutant LEC-6.

**Results**

**Overall structure of *C. elegans* galectin LEC-6**

We determined the crystal structure of *C. elegans* LEC-6 in complex with Gal\(\beta\)-1-4Fuc at 1.5 Å resolution, using the molecular replacement method with a model structure based on human galectin-7 (protein data bank (PDB) ID: 1BKZ) as a search model (Leonidas et al. 1998; DiMaio et al. 2011). There are two LEC-6 molecules in the asymmetric unit (Figure 1A). Residues excluding some N-terminal amino acids and the N-terminal FLAG tag were modeled into the electron density map (molecule A: 6-146 amino-acid residues of nontagged LEC-6, molecule B: 7-146 amino-acid residues of nontagged LEC-6). Two Gal\(\beta\)-1-4Fuc molecules and three magnesium ions were located in the difference Fourier map. The final crystallographic model was refined to an R-value of 14.7% and R\(_{free}\)-value of 17.9%. Further details of the structure determination are described in “Materials and methods” section. Statistics of the structure determination and its refinement are summarized in Supplementary data, Table S1. The crystals of LEC-6 belong to the space group P6_3; the three magnesium ions, derived from the crystallization conditions, contribute to the crystal packing. The electrostatic surface potential map of this LEC-6 dimer is represented in Figure 1B. The surface of the monomer-connecting region is more highly charged than the other regions of the LEC-6 dimer.

**Monomer structure and dimer interface**

Each LEC-6 monomer in the asymmetric unit has a typical galectin fold. Two antiparallel six-stranded \(\beta\)-sheets (F1–6 and S1–6) form a \(\beta\)-sandwich in the LEC-6 monomer (Figure 2A); Three continuous strands (S4, 5 and 6) form the binding site for Gal\(\beta\)-1-4Fuc. The most likely dimer interface encompasses the sides of F1 and S1 from each molecule derived by the program PDBePISA (Figures 1A and 2B).

Fig. 1. The overall structure of *C. elegans* galectin LEC-6. (A) The overall structure of the LEC-6 dimer (molecule A colored yellow and molecule B colored cyan and light pink) in complex with Gal\(\beta\)-1-4Fuc (colored green). Two orthogonal views of the dimer are shown. (B) Electrostatic potential maps of the dimer surfaces of LEC-6 (upper) from the same orientation as in the lower panel of A and the flip side (lower). Surface potentials range from −10 kT (red) to 10 kT (blue).
The buried surface area is 1909 Å², representing the sum of areas from both molecules. The complex formation significance score (CSS) of this interface is 0.974, and the shape complementarily statistics (Sc) value is 0.837. The architecture of dimer orientation is the same as that of human galectin-1 and -2, but opposite to that of human galectin-9 N-terminal CRD (Nagae et al. 2006). Each F1-strand interacts with the F1 and S1 strands of the other molecule in the dimer (Figure 2B and C). The residues Gly7, Ile8, Phe10, Asn12 and Pro24 in molecule A form hydrogen bonds with residues Gly7, Ser9, Phe10, Pro24 and Gln26 in molecule B (Figure 2C). Additional hydrogen bonds with water molecules, as well as hydrophobic interactions, are involved in the LEC-6 dimer interface. We confirmed the dimer formation in solution by performing a chemical cross-linking experiment (Supplementary data, Figure S1). We also performed dynamic light scattering (DLS) experiments to elucidate the solution state of LEC-6. The hydration radius and the estimated molecular weight were 2.92–2.96 nm and 41–43 kDa, respectively. The values are slightly larger than, but close to, the predicted values for a dimer calculated from our crystal structure (radius of gyration: 2.33 nm) and the amino-acid sequence (monomer MW is 16 kDa). DLS measurements sometimes show larger values because of the effect of hydration water molecules. Based on the results of this DLS experiment and the chemical cross-linking experiment, we conclude that LEC-6 forms a dimer in solution.

Fig. 2. The interface of the dimer structure of LEC-6. (A) Secondary elements and standard numbering are indicated on the LEC-6 monomer (“F1–6” colored cyan; “S1–6” colored light pink) in complex with Galβ1-4Fuc (colored green). (B) A Schematic representation of the LEC-6 dimer, with the same colors and standard numbering as in Figures 1A and 2A. The two hexagonal shapes with lines on each S4–S6 region represent Galβ1-4Fuc molecules (“G” means galactose moiety and “F” means fucose moiety). (C) The close-up view of the LEC-6 dimer interface, with the same colors and standard numbering as in Figure 2B. Residues that form the hydrogen bonds in the dimer interface are represented in stick representation.

Alternative conformation of fucose
The electron density corresponding to the oligosaccharide in molecule A of LEC-6 can be fitted into alternative conformations (Supplementary data, Figures S2 and S3). When the alternative conformation of Galβ1-4Fuc in LEC-6 was fitted to the positive-difference density from the refinement with the original position of Galβ1-4Fuc (Supplementary data, Figure S3B), the significant difference density appeared at the position of the alternative conformation of Galβ1-4Fuc (Supplementary data, Figure S3C). The original position of Galβ1-4Fuc provided a good fit to the difference density from the refinement with the alternative conformation of Galβ1-4Fuc (Supplementary data, Figure S3D). If there were no interactions between Glu67 of LEC-6 and the hydroxyl group of Galβ1-4Fuc, then there would be no significant positive-difference density at the original position of Galβ1-4Fuc after the refinement with the alternative conformation of Galβ1-4Fuc. There is no neighboring...
molecule in the vicinity of the Galβ1-4Fuc binding site of molecule A (Supplementary data, Figure S2A); however, a neighboring molecule is positioned close to the Galβ1-4Fuc binding site of molecule B (Supplementary data, Figure S2B). Crystal packing appeared to stabilize one conformation of Galβ1-4Fuc, but did not form an artificial conformation in molecule B. Unless otherwise noted, hereafter we refer to the structural details of molecule B.

Sugar-binding site in the LEC-6 structure

After molecular replacement and initial model building, there remained a large unmodeled density near S4, S5 and S6. The Galβ1-4Fuc disaccharide fits well to this residual electron density. Twelve hydrogen bonds contribute to the interaction between LEC-6 and Galβ1-4Fuc (Figure 3A). The residues His60, Asn62, Arg64, Glu67, Asn73, Glu83 and Arg85 are involved in the formation of hydrogen bonds to Galβ1-4Fuc. The Trp80 side chain stacks against the galactose moiety of Galβ1-4Fuc. Residues His60, Arg64, Asn73, Trp80 and Glu83 are hydrogen-bonded to the galactose moiety of the disaccharide. These residues are well conserved in other galectins, with the exception of LEC-7 (Figure 4). In contrast, the residues Glu67 and Arg85, which are only involved in binding to the fucose moiety, show variation between LEC-6 and other galectins (Figure 4). This difference seems to explain the differences among galectins with respect to the specificity of disaccharide binding.

Recently, the structure of C. cinerea galectin CGL2 in complex with Galβ1-4Fucα1-6GlcNAc was solved by X-ray crystallography (Butschi et al. 2010). A comparison of the carbohydrate-binding site with that of LEC-6 reveals that the residues involved in carbohydrate binding are conserved, with the exception of Asn62LEC-6 (Ser53CGL2) (Figure 3). Our previous paper showed that Asn62LEC-6 is not critical for binding of LEC-6 to Galβ1-4Fuc (Takeuchi et al. 2009). The arrangement of side chains from these conserved residues in LEC-6 and CGL2 could underlie the mechanism of the specificity of binding of the fucose moiety by both galectins.

FAC analysis

Based on the X-ray crystal structure of the complex between LEC-6 and Galβ1-4Fuc, we focused on the role of Glu67LEC-6 in recognition of the fucose moiety, because we have already examined the effect of the mutation of each of the other eight residues on sugar-binding ability (Takeuchi et al. 2009). To confirm the importance of this residue, we made recombinant LEC-6 wild-type and E67D and E67A point mutants, and performed FAC analyses to observe the change in affinity for saccharides. The obtained $K_a$ values are shown in Figure 5; The $K_a$ values of LEC-6 R85H mutant (Takeuchi et al. 2009) are shown for the comparison. Both E67D and E67A mutants exhibited weaker affinity for E3 in comparison with wild-type (Figure 5). The E67D mutation significantly reduced the ability to bind E3, a representative endogenous oligosaccharide ligand N-glycan chain of...
C. elegans glycoproteins containing Galβ1-4Fucα1-6GlcNAc structure: It showed no affinity for galactosidase-treated E3 (E3 + Galase), suggesting the specific recognition of the galactose residue. In contrast, the ability to bind sugars containing Galβ1-4GlcNAc units (NA2, NA3 and NA4) was not significantly affected. The E67A mutant showed a similar tendency, but its affinities for all the tested saccharides were lower than those of the E67D mutant. Alanine mutants often have severe effects, because the side chain is short and has no polarity. These results suggest the importance of Glu67 for the interaction with the fucose moiety. Furthermore, the E67D mutant also exhibited reduced affinity for a synthetic oligosaccharide, Galβ1-4Fuc-Man-ol-PA, in which 1-OH of fucose is linked via a mannitol-derived spacer to pyridylamine, though the effect was less significant; This tendency was also observed in the mutant of Arg85LEC-6, which also contributes to the interaction with the fucose moiety. This synthetic disaccharide derivative does not have a branched structure, and might therefore adapt more easily to the binding site of a mutant protein in which the glutamic acid residue (Glu67) has been replaced by aspartic acid.

**Discussion**

The solved three-dimensional structure of LEC-6 protein is similar to those of the N-terminal CRDs of human galectin-8 and -9 among the protein 3D-structure database including...
other galectins by Dali search (Holm and Rosenström 2010). The sequence identity of galectin-8 (N-terminal CRD: 1–148) and gelcettin-9 (N-terminal CRD: 1–154) against LEC-6 are 25.3 and 28.3%, respectively. Although LEC-6 and both of these human galectins contain the well-conserved residues for binding to disaccharide units, Glu67LEC-6 is not conserved in the N-terminal CRDs of the human galectins. Residues Arg72gal-8 and Asp68gal-9 correspond to Glu67LEC-6 in the multiple sequence alignment (Figure 4). We compared the crystal structure of LEC-6 to those of the N-terminal CRDs of galectin-8 and -9, whose structures have been determined in complex with lactose (PDB ID: 3AP4 and 2EAK) (Nagae et al. 2010). The electron densities corresponding to the oligosaccharide in the CGL2 crystal structure (Figure 3B). The electron densities corresponding to the oligosaccharide in the LEC-6 and CGL2 crystal structures can be fitted into alternative conformations (Supplementary data, Figures S3 and S5) (Butschi et al. 2010). Our FAC analyses of Glu67 and Arg85 mutants of LEC-6 also demonstrated that their affinities for E3 were more significantly reduced than their affinities for Galβ1-4Fuc (Figure 5) (Takeuchi et al. 2009). These data suggest that residues Glu67 and Arg85 are more important for the binding of endogenous oligosaccharides than for di- and tri-saccharides, not only in LEC-6, but also in CGL2. Among them, we conclude that the conserved Glu67 assists the preferential binding of fucose moieties.

**Materials and methods**

**Materials**

Galβ1-4Fuc and Galβ1-4Fuc labeled with pyridylamine via a spacer (Galβ1-4Fuc-Man-ol-PA) were synthesized chemically (Nishiyama et al. 2010). NA2-PA (PA-001, Galβ1-4GlcNAcβ1-2Manα1-3 (Galβ1-4GlcNAcβ1-2Manα1-6) Manβ1-4GlcNAcβ1-4GlcNAc-PA), NA3-PA (PA002, Galβ1-4GlcNAcβ1-2 (Galβ1-4GlcNAcβ1-4) Manα1-3 (Galβ1-4GlcNAcβ1-2Manα1-6) Manβ1-4GlcNAcβ1-4GlcNAc-PA), NA4-PA (PA004, Galβ1-4GlcNAcβ1-2 (Galβ1-4GlcNAcβ1-4) Manα1-3 (Galβ1-4GlcNAcβ1-2 (Galβ1-4GlcNAcβ1-4) Manα1-6) Manβ1-4GlcNAcβ1-4GlcNAc-PA), and rhamnose-PA were purchased from Takara Bio (Shiga, Japan). E3-PA, a pyridylaminated endogenous N-glycan ligand of LEC-6, galactosidase-treated E3-PA were prepared as reported previously (Takeuchi et al. 2008).

**Fig. 5.** FAC analysis of LEC-6 E67D and E67A mutants. (A) Structures of the PA-sugars used in FAC analysis. Open circle with diagonal line, hexose; open circle, mannose; filled circle, galactose; filled square, GlcNAc; open triangle, fucose. (B) Bar graph representation of the $K_i$ values for the interaction between LEC-6 and PA-sugars. The $K_i$ values for these interactions were calculated as described in “Materials and methods” section. R85H mutant data from our previous report (Takeuchi et al. 2009) are provided for comparison. Data are expressed as the mean ± SD ($n = 3$). N.D.: not determined.
Protein purification

Escherichia coli strain BL21(DE3) was transformed with the pET-FLAG-LEC-6 expression plasmid (Takeuchi et al. 2009), cultivated overnight at 37°C in Luria-Bertani medium supplemented with 50 μg/mL of ampicillin (LBA medium), transferred to a 25×volume of LBA medium, and incubated at 37°C for 3 h. After chilling the culture, isopropyl-1-thio-β-D-galactopyranoside was added to the culture at a final concentration of 0.2 mM, and the culture was further incubated overnight at 20°C. The cells were harvested, suspended in TBS-DTT (20 mM Tris–HCl, 150 mM NaCl, 1 mM dithiothreitol, pH 7.5), and lysed by sonication. After centrifugation, the supernatant was applied to an immobilized asialofetuin column prepared as described previously (Arata et al. 1997). After extensive washing of the column with TBS-DTT, the adsorbed protein was specifically eluted with TBS-DTT containing 0.1 M lactose. The buffer of the eluate was changed and the purified recombinant protein was concentrated by centrifugation using an Amicon Ultra-15 filter device (Millipore, Billerica, MA). About 10 mg of purified LEC-6 protein was obtained from 1 L coli culture. The purity was estimated to be >95% by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. The resulting protein solution was used for crystallization.

FAC analysis of LEC-6 mutant

Point mutants of LEC-6 were generated by PCR using the pGEM-LEC-6 plasmid (Takeuchi et al. 2009) and the following primers (substitution sites are underlined): E67D, 5′-ATGGAGCTTGTGTAACAATCT-3′ and 5′-CATCAGATTC TGGCGTTGAAGTG-3′; E67A, 5′-CGGGAAGCTTTGTA ACAAC-3′ and 5′-CATCGAATCTAGCGTTGAAGTG-3′. The resulting mutated LEC-6 gene was then subcloned into the pET-FLAG expression vector. Expression and purification of LEC-6 mutants were performed basically as described previously (Arata et al. 1997), except for the use of an immobilized Galβ1-4Fuc adsorbent (Takeuchi et al. 2011).

Immobilization of LEC-6 wild-type, E67D and E67A proteins on HiTrap N-hydroxysuccinimide-activated Sepharose (GE Healthcare, St. Giles, UK) and FAC analysis were performed essentially as described previously (Takeuchi et al. 2009). For FAC analysis, fluorescence-labeled oligosaccharides (PA-sugars) at a concentration of 5 mM were applied to a column containing an immobilized LEC-6 mutant (5.1 mg protein/mL gel), at a flow rate of 0.25 mL/min at 20°C. Elution of PA-sugar from the column was monitored using a fluorescence detector. The Kd values for the interactions between each LEC-6 mutant and PA-sugar were determined according to the following equation: 

\[ K_d = B_o (V_f - V_o) / [A]_0 \]

In this equation, \(B_o\) is the effective ligand content, \(V_f\) is the volume of the elution front, \(V_o\) is the \(V_f\) of rhamnose-PA (which is not bound by LEC-6 and used as a negative control), and \([A]_0\) is the initial concentration of the PA-sugar. In this study, the \(B_o\) value of each immobilized LEC-6 mutant column was calculated from the data obtained by concentration-dependence analysis using various concentrations of Galβ1-4Fuc. The \(K_a\) values were calculated based on the equation, \(K_a = 1/K_d\).

Crystallization and X-ray data collection

For structural analysis, the protein was concentrated at 10.6 mg/mL in 20 mM Tris–HCl pH 7.5, 50 mM NaCl and 10 mM Galβ1-4Fuc. The crystallization screening was performed by a crystallization robot (Hiraki et al. 2006). The crystals used for diffraction were obtained in 25% w/v PEG-3350, 100 mM HEPES pH 7.5 and 200 mM MgCl2. Diffraction data were collected at 100 K on beamline BL-17A of the Photon Factory, KEK at Tsukuba, JAPAN (Igarashi et al. 2007) and processed with HKL2000 (Otwinowski et al. 1997) Crystal parameters and data-collection statistics are summarized in Supplementary data, Table S1.

Structure determination and refinement

The LEC-6 crystal structure was solved by the molecular replacement method using Phaser (McCoy et al. 2007; Adams et al. 2010). The search model for molecular replacement (Q9N384) was created by ModBase (Pieper et al. 2011) in UniProt (Jain et al. 2009) from LEC-6 sequence and based on the human galectin-7 structure (PDB ID: 1BKZ). The amino-acid sequence identity is 28% between LEC-6 and human galectin-7. After molecular replacement, \(R\) and \(R_{free}\) values were 0.25 and 0.27, respectively. Since the values were reasonable but slightly high at 1.5 Å (highest resolution), density modification (Parrot in CCP4) (Collaborative Computational Project 1994) was applied to improve the phases from molecular replacement. The initial model was built using ARP/wARP in CCP4 (Langer et al. 2008). After building the initial model, \(R\) and \(R_{free}\) values reduced to 0.20 and 0.23, respectively, and the large unmodeled electron density, similar in shape to a disaccharide, appeared. Further manual protein and ligand building were performed using Coot (Emsley et al. 2010). The coordinate and topology files for Galβ1-4Fuc were obtained from the sweet2 server (Bohne et al. 1999) and the GlycoBioChem PRODRG2 server (Schuttelkopf and van Aalten 2004), respectively. The model refinement was carried out using REFMAC v5.6 (Murshudov et al. 1997) with the TLS option in the CCP4 package. The quality of the final model was validated using the MolProbity server (Chen et al. 2010). The data collection and refinement statistics are summarized in Supplementary data, Table S1. Figures showing structural representations were prepared using RasMol (Bernstein 2000), Molscript (Kraulis 1991) and Raster3D (Merritt and Bacon 1997). The electron density map and electrostatic potential were calculated using PyMOL (Schrodinger 2010) with the APBS package (Baker et al. 2001) including PDB2PQR (Dolinsky et al. 2007). The interface of the structures was analyzed using the PISA server (Krissinel and Henrick 2007) and the Sc program in the CCP4 package. The residues involved in binding between LEC-6 and the carbohydrate were assigned using the Contact program in the CCP4 package and Ligplot (Wallace et al. 1995). PDB coordinates are deposited (3VV1).

Dynamic light scattering

DLS experiments were carried out four times using DynaPro NanoStar (Wyatt Technology Co., Santa Barbara, CA) at a LEC-6 concentration of 5 mg/mL, which corresponds to the protein concentration in the initial crystal screening. The data
were analyzed by the DYNAMICS software (Wyatt Technology Co.).

Supplementary data
Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Acknowledgements
We thank Mr. T. Ohota for his technical support.

Conflict of interest
None declared.

Abbreviations
CGL2, Coprinopsis cinerea galectin-2; CRD, carbohydrate-recognition domains; CSS, complex formation significance score; DLS, dynamic light scattering; FAC, frontal affinity chromatography; Fuc, fucose; GALT-1, galactosyltransferase; Galβ1-4Fuc, galactose-β1,4-fucose; GloNAc, N-acetyl glucosamine; LBA medium, Luria-Bertani medium supplemented with 50 μg/mL of ampicillin; PDB, Protein Data Bank.

References


