

Directed Evolution of Lectins with Sugar-binding Specificity for 6-Sulfo-galactose^{*[5]}

Received for publication, February 9, 2012, and in revised form, April 3, 2012. Published, JBC Papers in Press, April 5, 2012, DOI 10.1074/jbc.M112.351965

Dan Hu, Hiroaki Tateno, Atsushi Kuno, Rikio Yabe, and Jun Hirabayashi¹

From the Research Center for Medical Glycoscience, National Institute of Advanced Industrial Science and Technology (AIST), Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan

Background: 6-sulfo-galactose is a potential marker of various diseases, but no useful probes to detect this glycoepitope have been available.

Results: Mutants with novel affinity for 6-sulfo-galactose were engineered from an R-type galactose-binding lectin.

Conclusion: We succeeded in providing a method to create probes for 6-sulfo-galactose by a molecular evolutionary strategy.

Significance: The tools developed in our study will be especially useful in the context of sulfoglycomics.

6-sulfo-galactose (6S-Gal) is a prevalent motif observed in highly sulfated keratan sulfate, which is closely associated with the glioblastoma malignancy while acting as a critical determinant for endogenous lectins. However, facile detection of this unique glycoepitope is greatly hampered because of a lack of appropriate probes. We have previously reported tailoring an α 2-6-linked sialic acid-binding lectin from a ricin-B chain-like galactose-binding protein, EW29Ch, by a reinforced ribosome display system following an error-prone PCR. In this study, we challenged the creation of novel lectins to recognize 6S-Gal-terminated glycans by incorporating a high-throughput screening system with a glycoconjugate microarray. After two rounds of selection procedures, 20 mutants were obtained and 12 were then successfully expressed in *Escherichia coli*, 8 of which showed a significant affinity for 6'-Sulfo-LN (6-O-sulfo-Gal β 1-4GlcNAc), which the parental EW29Ch lacked. Analysis of two representative mutants by frontal affinity chromatography revealed a substantial affinity ($K_d \sim 3 \mu\text{M}$) for a 6S-Gal-terminated glycan. On the basis of the observation that all eight mutants have a common mutation at Glu-20 to Lys, site-directed mutagenesis experiments were performed focusing on this aspect. The results clearly indicated that the E20K mutation is necessary and sufficient to acquire the specificity for 6S-Gal. We also confirmed a difference in binding between E20K and EW29Ch to CHO cells, in which enzymes to catalyze the synthesis of 6S-Gal were overexpressed. The results clearly demonstrate that these mutants have potential to distinguish between cells containing different amounts of 6S-Gal-terminated glycans. This new technology will be used to provide novel tools essential for sulfoglycomics.

6-sulfo-Gal (6S-Gal)² is a characteristic motif of highly sulfated keratan sulfate (KS), which is associated with many bio-

logical phenomena, such as the malignancy of glioblastoma and astrocytic tumors (1, 2). Synthesis of 6S-Gal is specially catalyzed by KS Gal 6-O-sulfotransferase (KSST, also denoted as GST1) through sulfation of GlcNAc-sulfated polylectosamine (3, 4). *In vitro* studies indicate that KSST also catalyzes the 6-O sulfation of Gal residues in sialyl *N*-acetylglucosamine but not in *N*-acetylglucosamine (5). Indeed, 6S-Gal has been found in *O*-linked glycans in both the L-selectin receptor glycoprotein GlyCAM-1 and MUC1 secreted from breast cancer cells (6, 7). To date, facile detection of this glycoepitope has been greatly hampered as no specific monoclonal antibodies have been available (8, 9). Although several endogenous lectins, including Langerin, Siglec-8, and Siglec-F, have been shown to recognize 6S-Gal-containing glycans (10–12), precise distribution of their ligand glycans *in vivo* and identification of their carrier proteins remain to be determined.

Use of exogenous lectins is an alternative approach to structural characterization of glycans, as has been shown in cell typing, histochemical staining, and glycoprotein fractionation (13–15). An advanced technique called lectin microarray, in which a collection of well defined lectins is immobilized on a solid support, has now been applied successfully for high-throughput glycome analysis targeting serum glycoproteins and even cells (14, 16, 17). However, currently available lectins, mostly derived from plants, have an apparent drawback in their “repertoire,” lacking some critical probes, such as those for sulfated glycans. Therefore, engineering lectins on the basis of existing ones toward novel sugar-binding activity would be of great practical value.

In directing lectin evolution, choosing an appropriate protein scaffold as a starting template is crucial. R-type lectins containing a ricin-B chain-like carbohydrate recognition domain are widely distributed in bacteria, plants, and animals. Notably, most of the R-type lectins bind Gal/GalNAc, but some have evolved to bind their derivatives, *e.g.* Neu5Ac α 2-6Gal/GalNAc (18, 19). These facts suggest that an R-type lectin is a good scaffold for the engineering of lectins. EW29Ch is the C-termi-

* This study was supported in part by the Japan Society for the Promotion of Science Scholarship.

[5] This article contains supplemental Figs. S1–S7.

¹ To whom correspondence should be addressed: Research Center for Stem Cell Engineering, National Institute of Advanced Industrial Science and Technology (AIST), Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan. Tel.: 81-29-861-3124; Fax: 81-29-861-3125; E-mail: jun-hirabayashi@aist.go.jp.

² The abbreviations used are: 6S-Gal, 6-sulfo-galactose; KS, keratan sulfate; KSST, KS Gal 6-O-sulfotransferase; PA, pyridylamino; Gn6ST, GlcNAc 6-O-

sulfotransferase; LN-PAA, *N*-acetyl-lactosamine-polyacrylamide; 6'-Sulfo-LN-PAA, 6'-sulfo-*N*-acetylglucosamine-polyacrylamide; FAC, frontal affinity chromatography.

Engineering Lectins Specific for 6-Sulfo-galactose

nal domain of earthworm 29-kDa Gal-binding protein (EW29) that contains two R-type domains in tandem (20). Recombinant EW29Ch is stable and can be easily produced in *Escherichia coli*, allowing it to be used for the purpose of engineering. In fact, we have previously tailored a novel $\alpha 2-6$ -linked sialic acid-binding lectin from EW29Ch by a procedure consisting of error-prone PCR and a reinforced ribosome display method (21).

In this study, we greatly improved the basic strategy by incorporating a high-throughput screening system, *i.e.* a glycoconjugate microarray on the basis of an evanescent field fluorescence-assisted detection principle (22), and challenged our system to engineer novel lectins to recognize 6S-Gal-terminated glycans. As a result, a series of mutant lectins showing a novel affinity for 6S-Gal-terminated glycans were selected. Most had a mutation at Glu-20 to Lys in common. Moreover, a single point mutation (E20K) proved to be sufficient to acquire this novel specificity. Finally, the E20K mutant was shown to be useful for *in vivo* experiments to detect the 6S-Gal glycoepitope in KSST-transfected cells.

EXPERIMENTAL PROCEDURES

Construction of Plasmids—pET27b-FLAG-EW29Ch containing an N-terminal FLAG tag before the EW29Ch coding sequence and pET27b-FLAG-EW29Ch-bio containing a C-terminal biotinylation sequence after FLAG-EW29Ch were prepared as described in detail in the supplemental information.

Selection by Ribosome Display—Generation of an EW29Ch random mutagenesis library was carried out as described previously (21) and is described in detail in the supplemental information. Selection of 6S-Gal binding mutants using biotinylated multivalent carbohydrate polymers by ribosome display is described in detail in the supplemental information. Briefly, *in vitro* translation of 2 μ g of EW29Ch library mRNA using the *E. coli* S30 extract system as a linear template (Promega) was incubated with biotinylated multivalent carbohydrate polymers (Glycotect) coupled on streptavidin-coated magnetic dynabeads (Invitrogen) at 4 °C for 1 h. In some cases, preblocking with excessive lactose (10 mM) was carried out before selection by carbohydrate polymer-coated beads. After the selection, bound mRNA was purified, reverse-transcribed, and amplified with a one-step RNA PCR kit (Takara). The derived cDNA fragment was ready for the next round of selection and cloning.

Analysis of Selected Mutants by Glycoconjugate Microarray—The sugar-binding activity of selected EW29Ch mutants was analyzed by glycoconjugate microarray on the basis of methods described previously (22) and as described in detail in the supplemental information. Briefly, EW29Ch mutants with an N-terminal FLAG tag in pET27b were expressed in BL21-CodonPlus (DE3)-RIL and lysed. Labeling of cell lysate was performed by incubation with anti-FLAG M2 antibody and Cy3-labeled goat anti-mouse antibody. The labeled solution was applied to a glycoconjugate microarray and detected by an evanescent-type scanner, Glycostation Reader 1200 (GP Biosciences Ltd., Yokohama, Japan) in Cy3 mode.

Purification of Recombinant Proteins—Purification of recombinant proteins is described in detail in the supplemental information. Briefly, N-terminal FLAG-tagged wild-type EW29Ch or

EW29Ch mutants, in some cases containing a C-terminal biotinylation sequence, were expressed in BL21-CodonPlus (DE3)-RIL and purified using a lactosyl-Sepharose column. Biotinylation of purified proteins with a C-terminal biotinylation sequence was carried out as described previously using biotin ligase BirA (23).

Site-directed Mutagenesis—Site-directed mutagenesis was performed using the Accuprime TaqDNA polymerase (Invitrogen) following the protocol of the manufacturer. Mutations were confirmed by DNA sequencing. The primers used are described in the supplemental information.

Frontal Affinity Chromatography (FAC)—FAC was performed as described previously (24, 25) and is described in detail in the supplemental information. Briefly, EW29Ch or its mutant recombinant proteins were immobilized on *N*-hydroxysuccinimide-activated Sepharose Fast Flow (GE Healthcare) and packed into a miniature column (inner diameter, 2 mm; length, 10 mm; bed volume, 31.4 μ l; Shimadzu). A panel of pyridylamino (PA)-labeled sulfated glycans was successively injected into the columns. The elution front of each PA glycan relative to that of the control (Man₅GlcNAc₂-PA), referred to as V-V₀, was then determined. The dissociation constant K_d was then calculated as described previously (24).

Flow Cytometry—CHO cells were maintained in RPMI1640 supplemented with 5% FBS, 100 units/ml penicillin, and 100 μ g/ml streptomycin. Transfection of CHO cells with KSST or GlcNAc 6-*O*-sulfotransferase (Gn6ST) or cotransfection with KSST and Gn6ST was performed using Lipofectamine LTX (Invitrogen). After 24 h, cells were harvested, and the binding of 5D4, wild type EW29Ch, and E20K was analyzed by flow cytometry. Briefly, 1×10^5 cells were incubated with 5D4 (diluted 500-fold), 10 μ g/ml biotinylated EW29Ch, or E20K on ice for 20 min followed by 2 μ g/ml Dylight 488-labeled goat anti-mouse antibody (Jackson ImmunoResearch) and 10 μ g/ml Cy3-labeled streptavidin (Jackson ImmunoResearch). The binding of lectins to cells was analyzed by FACSCanto-II cytometer (BD Biosciences). For tunicamycin or swainsonine treatment, KSST-transfected cells were incubated with 1 μ g/ml tunicamycin and 5 μ g/ml swainsonine for 12 h before analysis, respectively. For sialidase treatment, cells were incubated with 25 milliunits of *Arthrobacter ureafaciens* sialidase at 37 °C for 30 min.

RESULTS

Refining the Strategy for Lectin Engineering—In our previous study, a ribosome display method along an error-prone PCR was used to direct evolution of EW29Ch for $\alpha 2-6$ -sialic acid-binding lectin. Selection of the target sialic acid-binding mutants was carried out using fetuin-coupled Sepharose beads as this glycoprotein is rich in terminal $\alpha 2-6$ sialic acid. To date, no glycoproteins that carry a mass of 6S-Gal-terminated glycans have been available, and even if there are any, they must be highly heterogeneous in a glycan structure. There is also a possibility that glycoproteins interact with the ribosome complex in their protein backbone, which would generate a “false positive” during selection. In this study, we used commercially available biotinylated carbohydrate polymers with a defined, homogeneous glycan structure as bait ligands for selection. To

test the feasibility of these carbohydrate polymers, we first compared the selection of wild-type EW29Ch by the target glycan *N*-acetyl-lactosamine-polyacrylamide (LN-PAA) and the non-target glycan mannose-polyacrylamide. As shown in supplemental Fig. S1A, LN-PAA-coated beads precipitated EW29Ch (in the form of a complex with mRNA) in a manner dependent on the amount of immobilized glycans. On the other hand, the nonspecific binding caused by the intact streptavidin beads themselves was gradually reduced as the increment of immobilized non-target mannose-polyacrylamide. It is highly possible that the hydrophilic carbohydrate polymers on the surface of the beads could inhibit the nonspecific interactions between ribosome complex and streptavidin-coated beads. At a concentration of 100 $\mu\text{g}/\text{ml}$, mannose-polyacrylamide beads did not substantially precipitate EW29Ch, whereas LN-PAA did so maximally. These results demonstrated the good selectivity of the carbohydrate polymers for the selection of lectins.

Having established suitable selection conditions, we attempted high-throughput screening of the desired mutants. We had previously used FAC, a quantitative method for determination of affinity constants (K_d) between lectins and glycans. However, this approach requires relatively large amounts of “purified” proteins and is time-consuming. In this study, we adopted a more advanced glycoconjugate microarray system that is based on a unique, evanescent field-assisted fluorescence detection principle (22). The method is directly applicable to crude extracts containing glycan-binding proteins with the aid of fluorescence-labeled specific antibodies. To examine the suitability of this method, wild-type EW29Ch fused with an *N*-terminal FLAG tag was expressed in *E. coli*, and the cells containing this protein were lysed. The expression of the target protein was visualized by SDS-PAGE, and the protein with a FLAG tag was confirmed by Western blotting using an anti-FLAG M2 antibody (supplemental Fig. S1B). Further, the cell extract was probed by incubation with anti-FLAG M2 antibody followed by Cy3-labeled secondary antibody. An aliquot of the labeled complex (EW29Ch/anti-FLAG M2 antibody/Cy3-labeled secondary antibody) was applied to the array, and the binding was monitored by an evanescent-type scanner at equilibrium without washing steps. As shown in supplemental Fig. S1C, clear binding signals were observed in FLAG-EW29Ch-expressed cells but not in a negative control (without isopropyl 1-thio- β -D-galactopyranoside induction), suggesting that the binding is specific to FLAG-EW29Ch. Notably, the binding profile was almost the same as that of the purified EW29Ch. All these data indicate that the glycoconjugate microarray should provide a better means to analyze the sugar-binding specificity of the selected mutants in a rapid, sensitive, and high-throughput manner.

Selection of 6S-Gal Binding Lectins—Having refined the engineering strategy, we performed random mutagenesis targeting the open reading frame of EW29Ch. Error-prone PCR was conducted under conditions where approximately six mutations per gene were introduced. The resultant mutant library was then selected against a sulfated carbohydrate polymer, 6'-sulfo-*N*-acetyl-lactosamine-polyacrylamide (6'-Sulfo-LN-PAA). Although the wild-type EW29Ch has an almost undetectable level of affinity to 6S-Gal-terminated glycans (described later), 6'-Sulfo-LN-PAA still precipitated a considerable amount of wild-type

and wild-type-like (neutral) mutants, possibly because they still have a significantly enhanced affinity toward multivalent carbohydrate polymers. To reduce this undesired interaction, we added excessive lactose to the selection system before mixing with 6'-Sulfo-LN-PAA-coated beads. On the basis of this concept, we first incubated the mutant library with excessive lactose (10 mM) for 1 h and then with 6'-Sulfo-LN-coated beads. As a result, the enriched population in the presence of lactose showed a large reduction in the RT-PCR product (background noise) after the first round of selection, probably because of bulk elimination of the undesired background binding of the wild type and its neutral mutants (Fig. 1A, upper panel). The obtained fraction was subjected to another round of ribosome display selection against 6'-Sulfo-LN. The enriched population no longer showed a substantial difference in the RT-PCR analysis regardless of the presence or absence of lactose (Fig. 1A, lower panel). This observation strongly suggests that some mutants with target specificity (6'-Sulfo-LN) were enriched.

To confirm this, isolated mRNA after two rounds of selection was reverse-transcribed, and the derived cDNA was treated with *Nde*I and *Xho*I to insert the coding fragment into pET27b containing an N-terminal FLAG tag. Twenty clones randomly selected were sequenced. As a common feature, 14 clones (1–14) were found to carry the mutation E20K in common in the α subdomain, and six had E68K in the β subdomain (Fig. 1B). These mutants were expressed in *E. coli*, and 12 were successfully expressed (1–8 and 15–18). The sugar-binding activity of the 12 mutants was analyzed by a glycoconjugate microarray consisting of eight related sugars (results of the wild-type and 4 are shown in Fig. 1C). Their relative affinity for 6'-Sulfo-LN was calculated by normalizing the signal with LN (Fig. 1D). Eight of the 12 mutants (1–8) carrying the mutation E20K exhibited a specific affinity for 6'-Sulfo-LN, whereas the rest of the clones (15–18) without this mutation lacked such affinity. Therefore, the point mutation E20K was strongly suggested as being important for the binding with 6'-Sulfo-LN.

Quantitative Analysis of the Affinity of the Selected Mutants to 6S-Gal-terminated Glycans by FAC—To assess the sugar-binding affinity of the derived mutants from a more quantitative aspect, two representative mutants (1 and 4) were purified to apparent homogeneity in SDS-PAGE and were subjected to FAC analysis as described previously (Fig. 2 and supplemental Fig. S2). Both mutants as well as the wild-type EW29Ch displayed a similar affinity to Gal-terminated glycans (901 and 902), *i.e.* in terms of K_d , 2×10^{-5} M. However, the two mutants exhibited a greatly increased affinity to a 6-sulfo-Gal-terminated glycan (922) ($K_d = 3\sim 4 \times 10^{-6}$ M) to which the wild-type EW29Ch showed no substantial affinity. Notably, no detectable affinity was shown to the related structures, including 3-sulfo-Gal (918), Neu5Ac α 2-6-Gal (704), and 6-sulfo-GlcNAc (924). A greatly reduced affinity to 923, the non-reducing terminally unsulfated structure of 922, unambiguously confirms the importance of 6-sulfation of Gal in the lacto-*N*-neotetraose structure (902). Comparison of the affinity toward a panel of 129 PA-labeled glycans (supplemental Fig. S3) revealed that EW29Ch and 4 have almost the same binding profiles except for terminal 6S-Gal glycan (922) (supplemental Fig. S2). These

Engineering Lectins Specific for 6-Sulfo-galactose

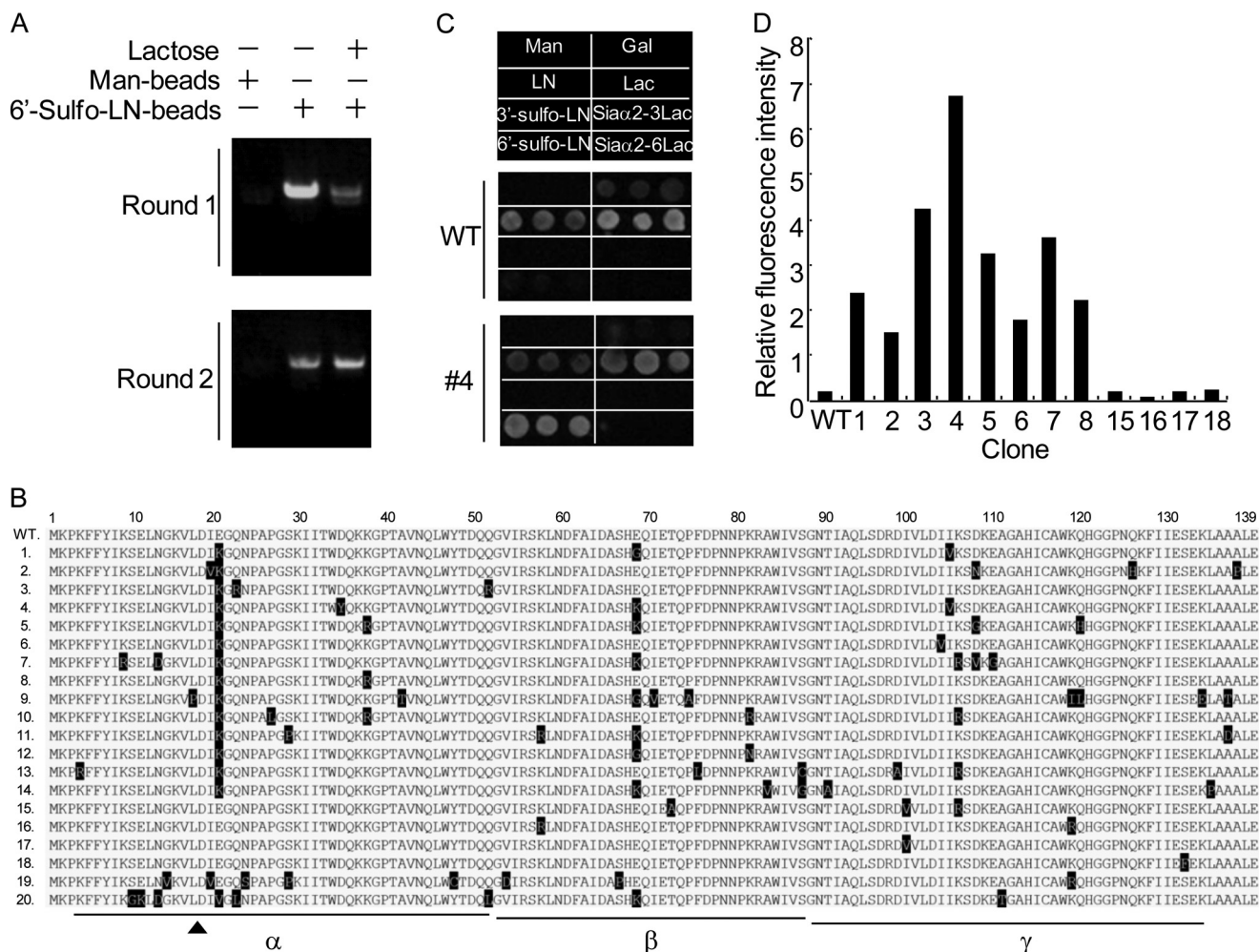


FIGURE 1. Selection of 6S-Gal binding lectins. *A*, selection of 6'-Sulfo-LN binding lectins by pretreatment of the translation mixture with excessive lactose. The EW29Ch library generated by error-prone PCR was selected with 6'-Sulfo-LN-coated beads with or without pretreatment of the translation mixture with lactose. The precipitated mRNAs under individual conditions were isolated and reverse-transcribed to cDNA by one-step RT-PCR. A part of cDNA was used for analysis by agarose electrophoresis. The rest of the cDNAs selected in the presence of lactose were subsequently used for the second round of ribosome display. In all experiments, mannose-polyacrylamide was used as a negative control. *B*, alignment of the amino acid sequences of 20 selected mutants along with the WT. The mutated amino acids are highlighted on *black background*. Three subdomains of EW29Ch, α , β , and γ , are *underlined*. The key residue Asp-18 in the binding with Gal is indicated by *arrows*. *C* and *D*, analysis of the sugar-binding activity of selected mutants by glycoconjugate microarray. Twelve mutants with an N-terminal FLAG tag (1–8 and 15–18) were expressed in *E. coli* and lysed. The resulting lysates were Cy3-labeled and then applied to a glass slide where eight kinds of glycans were immobilized. The binding of mutants to individual glycans was monitored by an evanescent-type scanner. *C*, the binding profile of wild-type EW29Ch and a representative mutant (#4). The relative signal of 6'-Sulfo-LN to LN of 12 expressing mutants is indicated in *D*.

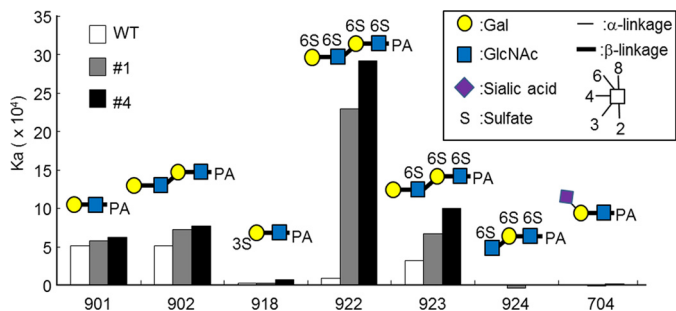


FIGURE 2. Quantitative analysis of the affinity of selected mutants to 6S-Gal-terminated glycans by FAC.

results show that the mutants (1 and 4) have acquired a novel affinity to 6S-Gal.

Investigation of the Effects of Hot Spot Mutation on the Affinity for 6S-Gal-terminated Glycans—Although the above results indicate that E20K (14 of 20) plays a critical role for the 6S-Gal

binding, the exact contribution of E20K is not clear, especially as another mutation E68K (5 of 14) often accompanied the E20K mutation. Mutants E20K, E68K, and the double mutant E20K/E68K were constructed by site-directed mutagenesis, and their sugar-binding specificities were analyzed by FAC (Fig. 3). As expected, the binding profiles of E20K were quite similar to those of the two mutants, 1 and 4, described above (compare with Fig. 2). E20K exhibited a much higher affinity to a 6S-Gal-terminated glycan (922) compared with a corresponding non-sulfated glycan (923). On the other hand, binding of E68K to the 6S-Gal-terminated glycan was not observed, and the double mutant E20K/E68K showed a similar profile to that of E20K. These results indicate that E68K is not essential for 6S-Gal recognition. The results also prompted us to investigate whether mutation of Glu-20 to other basic residues could acquire the target activity. For this, E20R and E20H were prepared, and their sugar-binding activities were analyzed by glycoconjugate

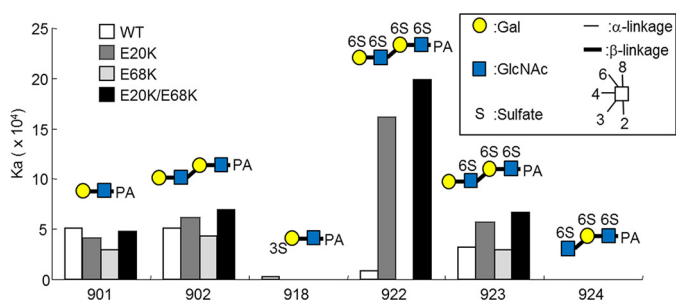


FIGURE 3. Quantitative analysis of the affinity of site-directed mutants to 6S-Gal-terminated glycans by FAC.

microarray. As shown in supplemental Fig. S4, E20R exhibited a comparable binding ability for 6'-Sulfo-LN to E20K, whereas E20H only slightly enhanced the affinity to 6'-Sulfo-LN. Taken together, these results show that the presence of a strong basic residue (Lys or Arg) at position 20 is necessary and sufficient for the acquisition of a novel affinity to 6S-Gal.

Binding of Mutants to CHO Cells in Which KSST or Gn6ST Were Overexpressed—To determine whether these mutants are of practical use for detecting 6S-Gal-terminated glycans in complex samples, CHO cells were transfected with a 6S-Gal catalyzing enzyme, KSST. Because KSST preferentially adds a sulfate group to Gal adjacent to a sulfated GlcNAc (5), cotransfection of KSST with Gn6ST was also performed. Cotransfection of KSST with Gn6ST not only enhances the expression of 6S-Gal but also that of highly sulfated KS that can be recognized by a specific antibody, 5D4 (26). Thus, binding of 5D4 was detected in KSST and Gn6ST cotransfected cells, suggesting that both enzymes are catalytically active when overexpressed in CHO cells (Fig. 4A). Both E20K and wild-type EW29Ch exhibited a similar binding to mock-transfected cells. However, when KSST, but not Gn6ST, was overexpressed, a stronger binding of E20K than EW29Ch was observed. The increased binding of E20K was further enhanced when KSST and Gn6ST were cotransfected (Fig. 4B). These results were also confirmed by staining of KSST and Gn6ST cotransfected cells (supplemental Fig. S5). E20K exhibited more intensive staining than that of EW29Ch. Double-staining of cotransfected cells with 5D4 and EW29Ch or E20K indicated that E20K, but not EW29Ch, preferentially bound to KS-overexpressed cells, suggesting that the enhanced binding of E20K was due to the expression of KS (Fig. 4C).

Considering that much of the utility of the mutants developed in this study should be in the detection of 6S-Gal on *N*- or *O*-linked glycans, not only for those on KS, to which specific antibodies (5D4) are already available, we are interested in the binding of E20K to KSST-transfected cells, to which binding of 5D4 was not observed (Fig. 4, A and B). To analyze whether these glycoepitopes are from the *N*-linked glycans, we treated the KSST-transfected cells with tunicamycin, a drug that inhibits the biosynthesis of all *N*-glycans, and swainsonine, an inhibitor of mannosidase II that can inhibit the biosynthesis of complex *N*-linked glycans. The effects of both inhibitors on the glycans of the cell surface were confirmed by binding with lectins (supplemental Fig. S6). Tunicamycin inhibited the binding of both ConA and *Phaseolus vulgaris* leucoagglutinin, whereas swainsonine showed an increased binding of ConA and

decreased *P. vulgaris* leucoagglutinin binding. Concomitant with the decrease of complex-type *N*-linked glycans on the cell surface, the stronger binding of E20K than EW29Ch to KSST-transfected cells was completely abolished by tunicamycin and partially inhibited by swainsonine (Fig. 4D), suggesting that most of 6S-Gal in KSST-transfected cells were possibly from the *N*-linked glycans. These results indicate that the mutants developed in this study are useful not only for the detection of 6S-Gal-bearing KS but also 6S-Gal-bearing *N*- or *O*-linked glycans.

Binding of Mutants to Sialidase-treated Cells—Although the newly developed lectins in this study have acquired the novel affinity to 6S-Gal, they still retain the affinity of the parental lectin EW29Ch to Gal. This seems to pose a significant problem when these lectins are used for detection of 6S-Gal in the samples, which are rich in non-sulfated Gal. To test the possibility, we investigated the binding of E20K and EW29Ch to the CHO cells that are treated with sialidase to increase the Gal on the cell surface (Fig. 5A). Although the binding of both E20K and EW29Ch to the Mock or KSST and Gn6ST cotransfected cells were significantly increased by the treatment of sialidase, the preferential binding of E20K than EW29Ch to KSST and Gn6ST cotransfected cells, in which KS-expressing cells comprised only 7% of the total cells (indicated by 5D4), was still observed. Therefore, it is expected that these lectins are useful for the detection of 6S-Gal even in the presence of a large amount of non-sulfated Gal.

To examine further whether these mutants could be used for the detection of endogenous 6S-Gal on cells other than the engineered CHO cells, we investigated the binding of E20K and EW29Ch to a glioblastoma cell line, LN229, which has been shown to express highly sulfated KS (2). As shown in Fig. 5B, none of these lectins showed significant binding to intact LN229 cells. However, we found a relatively stronger binding of E20K than EW29Ch after the treatment with sialidase, suggesting that these cells contain sialylated 6S-Gal on the cells, although further structural analysis is necessary for this proof.

Taken together, the above results indicate that these mutants are potentially useful tools in distinguishing cells containing 6S-Gal-bearing glycans.

DISCUSSION

In this study, lectins with novel affinity to 6S-Gal-terminated glycans were obtained from error-prone PCR-based lectin library after only two rounds of the ribosome display selection. Although the essence of the method has already been established (27), the high efficiency of enrichment achieved in this study is attributed to the several modifications made to the previous strategy. First, reduction of nonspecific (background) selection was achieved by use of synthetic biotinylated carbohydrate PAA as glycan ligands instead of natural glycoproteins, which are obviously heterogeneous in the glycan moieties. When these biotinylated carbohydrate polymers were coupled on the streptavidin-coated magnetic beads at high density (concentration 100 $\mu\text{g/ml}$), they could not only capture the mutant with target activity in the solution but also reduce the nonspecific interaction caused by the streptavidin-coated magnetic beads (supplemental Fig. S1A). Second, specific enrichment of

Engineering Lectins Specific for 6-Sulfo-galactose

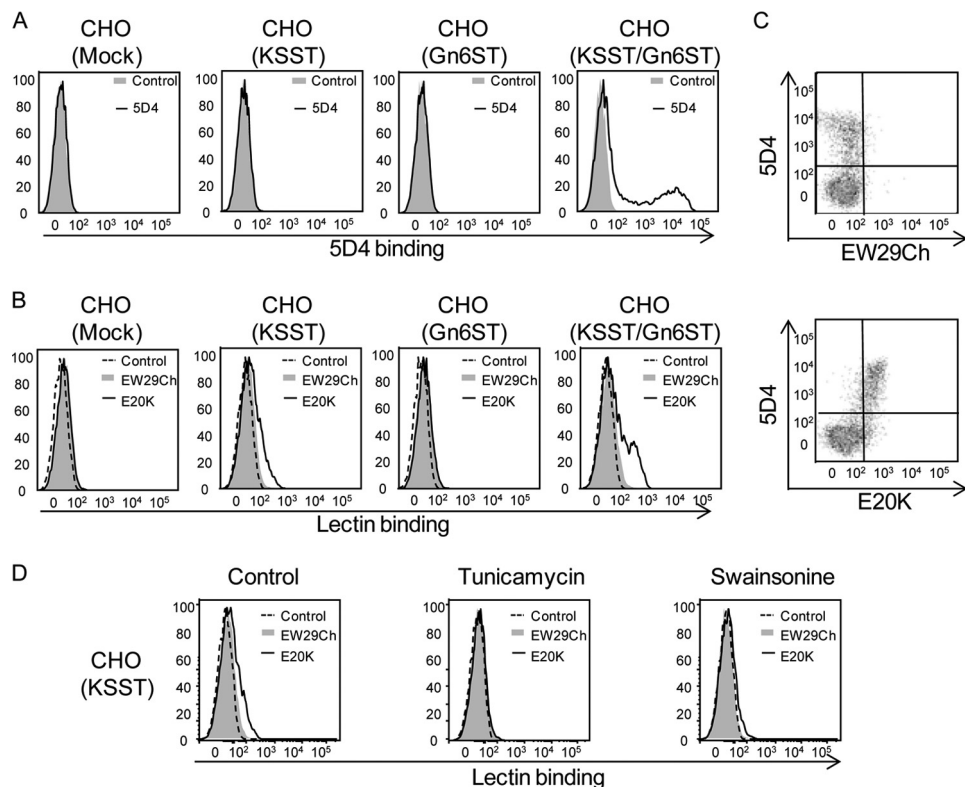


FIGURE 4. Binding of mutants to CHO cells in which 6S-Gal-catalyzing enzymes were overexpressed. The CHO cells were transfected with KSST or Gn6ST or cotransfected with KSST and Gn6ST for 24 h. The binding of 5D4 (A), wild-type EW29Ch, and E20K (B) to these transfected cells was investigated by flow cytometry as described under "Experimental Procedures." C, the double staining of KSST and Gn6ST cotransfected cells with 5D4 and EW29Ch or E20K. D, the binding of EW29Ch and E20K to KSST-transfected cells that were treated with 1 μ g/ml tunicamycin or 5 μ g/ml swainsonine for 12 h.

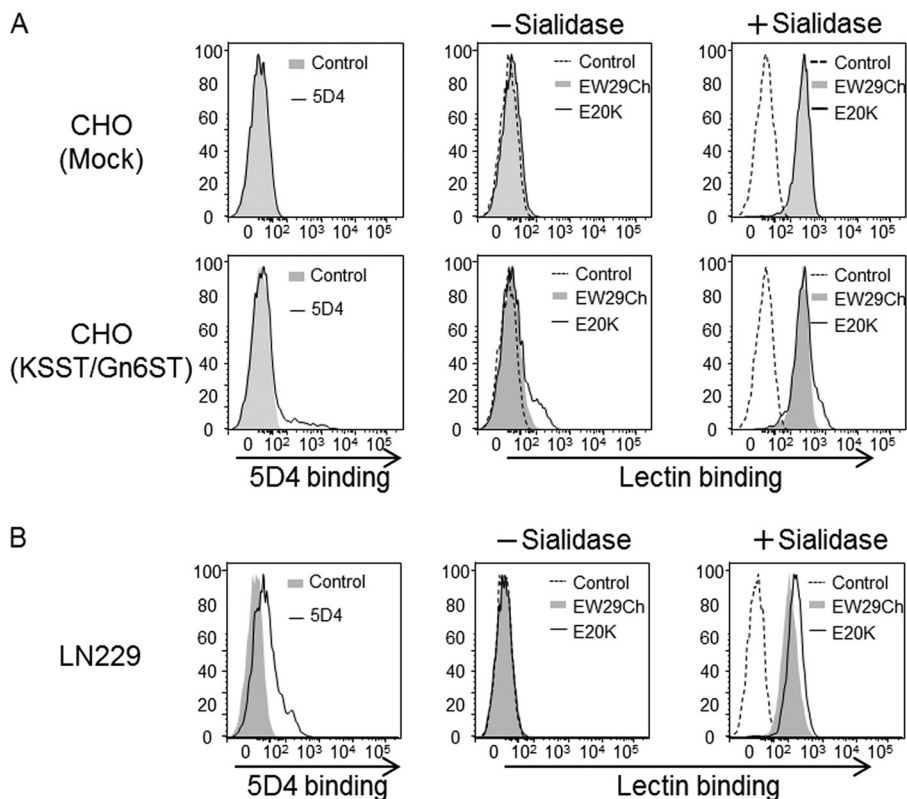


FIGURE 5. Binding of mutants to sialidase-treated cells. A, the binding of 5D4, EW29Ch, and E20K to the mock or KSST and Gn6ST cotransfected cells with or without sialidase treatment is shown. B, the binding of 5D4, EW29Ch, and E20K to a glioblastoma cell line, LN229, with or without sialidase treatment is shown.

target mutants was achieved by preblocking with excessive lactose to prevent the undesired selection of wild types or those preserving similar activity (neutral mutants).

In addition, we have introduced the advanced technology of evanescent field fluorescence-assisted glycoconjugate microarray into our evaluation system in place of FAC to facilitate screening lines of mutants derived by the above selection procedures. Because the microarray technique provides us with a highly sensitive and high-throughput screening, it greatly facilitated analysis of the sugar-binding activity of a series of candidate clones even in the form of crude extracts (supplemental Fig. S1C).

With the improved method, we took up the challenge of engineering lectins having specificity to 6S-Gal-bearing glycans because no effective probes to this biologically important glycoepitope have been reported. TJA-I is a quite interesting lectin purified from tuberous roots of *Trichosanthes japonica*, showing binding to 6S-Gal. However, it can also recognize both Neu5Ac α 2-6Gal and terminal Gal (28). Therefore, this property will hinder its use for specific detection of 6S-Gal. On the other hand, the 6S-Gal binding lectins developed in the present study exhibited substantially no binding to sialylated glycans (supplemental Fig. S2). Though these mutants still retain the affinity of the parental lectin to Gal, specific detection of 6S-Gal on cells can be achieved by combinational use with EW29Ch. Because E20K and EW29Ch have almost the same binding profile except for 6S-Gal (Fig. 3), thus, if we see increased binding of E20K to cells compared with EW29Ch, it is highly possible that these cells contain 6S-Gal. Consistently, we observed a stronger binding of E20K than EW29Ch to KSST-transfected cells, or KSST and Gn6ST cotransfected cells, in which 6S-Gal was overexpressed (Fig. 4B). Moreover, the preference was maintained even in the presence of a large amount of non-sulfated Gal after the sialidase treatment (Fig. 5). Thus, the 6S-Gal binding lectins developed in this study will be of great use in detection of the unique glycoepitope 6S-Gal.

In our previous study (21), we have chosen EW29Ch as a starting template because of several properties advantageous for lectin engineering. They include 1) adequate size (~14.5 kDa) for stability and error-prone PCR; 2) complete solubility, enabling high productivity in a conventional *E. coli* expression system; and 3) knowledge of the protein structure on the basis of both x-ray crystallography (29) and NMR analysis (30). A previous x-ray crystallographic analysis revealed that EW29Ch has a ricin-B chain-like β -trefoil structure consisting of three homologous (α , β , and γ) subdomains, with the α and γ subdomains possessing the Gal-binding activity (29). Later study by NMR titration experiments demonstrated that the α subdomain predominantly contributes to the Gal-binding activity of EW29Ch (30). Multiple hydroxyl groups attached to the non-reducing terminal Gal, except for the C6 position, have been shown to be directly involved in the hydrogen bonding with the hydrophilic residues Asp-18, Lys-36, and Asn-41 in the α subdomain (supplemental Fig. S7A). We therefore considered EW29Ch as the optimal template to direct evolution targeting 6-O-modified Gal (e.g. Neu5Ac α 2-6Gal, 6S-Gal) with the least destruction of the global structure. As a result, we observed "convergent" evolution at a molecular level. An acidic amino

acid, Glu at position 20, was found to be changed to basic Lys in 70% of the selected clones for acquisition of a novel binding activity for 6S-Gal, whereas the parental EW29Ch has no such activity. Glu-20 located two amino acids downstream of invariant Asp-18, a key residue in the binding with Gal (supplemental Fig. S7A). In the reported crystal structure of EW29Ch, Glu-20 is located in a loop structure proximal to the 6-OH of Gal. Therefore, it is anticipated that an electrostatic interaction between the negative sulfo- group of 6S-Gal and the positive Lys-20 of the mutant would be formed, which could account for the enhanced activity for 6S-Gal-terminated glycans. We also observed that changing Glu to another basic amino acid, Arg, improved the binding activity for 6S-Gal-terminated glycans, whereas His had only a limited effect (supplemental Fig. S4). Preferred enrichment of E20K rather than E20R can be attributed, at least in part, to single-nucleotide mutagenesis (GAR to AAR) to cause this code change, rather than dual mutagenesis (GAR to CGN/AGR). Interestingly, sequence alignment of subdomains from EW29Ch and several other R-type lectins indicates that mistletoe lectin-1 (ML1), ricin, and its non-toxic isolectin RCA120, possess a basic residue (Arg) in the corresponding position (supplemental Fig. S7B). This may explain why RCA120 bound strongly to 6S-Gal glycans in a recent study (31).

Overall, we have succeeded here in engineering novel lectins for 6S-Gal-terminated glycans for which no practical probes have been available. Thus, these lectins should be highly useful for future studies involving sulfated glycans, *i.e.* sulfoglycomics.

REFERENCES

- Kato, Y., Hayatsu, N., Kaneko, M. K., Ogasawara, S., Hamano, T., Takahashi, S., Nishikawa, R., Matsutani, M., Mishima, K., and Narimatsu, H. (2008) Increased expression of highly sulfated keratan sulfate synthesized in malignant astrocytic tumors. *Biochem. Biophys. Res. Commun.* **369**, 1041–1046
- Hayatsu, N., Ogasawara, S., Kaneko, M. K., Kato, Y., and Narimatsu, H. (2008) Expression of highly sulfated keratan sulfate synthesized in human glioblastoma cells. *Biochem. Biophys. Res. Commun.* **368**, 217–222
- Fukuta, M., Inazawa, J., Torii, T., Tsuzuki, K., Shimada, E., and Habuchi, O. (1997) Molecular cloning and characterization of human keratan sulfate Gal-6-sulfotransferase. *J. Biol. Chem.* **272**, 32321–32328
- Akama, T. O., Misra, A. K., Hindsgaul, O., and Fukuda, M. N. (2002) Enzymatic synthesis *in vitro* of the disulfated disaccharide unit of corneal keratan sulfate. *J. Biol. Chem.* **277**, 42505–42513
- Torii, T., Fukuta, M., and Habuchi, O. (2000) Sulfation of sialyl N-acetylglucosamine oligosaccharides and fetuin oligosaccharides by keratan sulfate Gal-6-sulfotransferase. *Glycobiology* **10**, 203–211
- Hemmerich, S., Leffler, H., and Rosen, S. D. (1995) Structure of the O-glycans in GlyCAM-1, an endothelial-derived ligand for L-selectin. *J. Biol. Chem.* **270**, 12035–12047
- Seko, A., Ohkura, T., Ideo, H., and Yamashita, K. (2011) Novel O-linked glycans containing 6'-sulfo-Gal/GalNAc of MUC1 secreted from human breast cancer YMB5 cells. Possible carbohydrate epitopes of KL-6(MUC1) monoclonal antibody. *Glycobiology* **22**, 181–195
- Kannagi, R., Ohmori, K., and Kimura, N. (2009) Anti-oligosaccharide antibodies as tools for studying sulfated sialoglycoconjugate ligands for siglecs and selectins. *Glycoconj. J.* **26**, 923–928
- Mitsuoka, C., Sawada-Kasugai, M., Ando-Furui, K., Izawa, M., Nakanishi, H., Nakamura, S., Ishida, H., Kiso, M., and Kannagi, R. (1998) Identification of a major carbohydrate capping group of the L-selectin ligand on high endothelial venules in human lymph nodes as 6-sulfo sialyl Lewis X. *J. Biol. Chem.* **273**, 11225–11233
- Tateno, H., Ohnishi, K., Yabe, R., Hayatsu, N., Sato, T., Takeya, M., Nari-

- matsu, H., and Hirabayashi, J. (2010) Dual specificity of Langerin to sulfated and mannosylated glycans via a single C-type carbohydrate recognition domain. *J. Biol. Chem.* **285**, 6390–6400
11. Bochner, B. S., Alvarez, R. A., Mehta, P., Bovin, N. V., Blixt, O., White, J. R., and Schnaar, R. L. (2005) Glycan array screening reveals a candidate ligand for Siglec-8. *J. Biol. Chem.* **280**, 4307–4312
 12. Tateno, H., Crocker, P. R., and Paulson, J. C. (2005) Mouse Siglec-F and human Siglec-8 are functionally convergent paralogs that are selectively expressed on eosinophils and recognize 6'-sulfo-sialyl Lewis X as a preferred glycan ligand. *Glycobiology* **15**, 1125–1135
 13. Fukushima, K., Satoh, T., Baba, S., and Yamashita, K. (2010) alpha1,2-Fucosylated and beta-N-acetylgalactosaminylated prostate-specific antigen as an efficient marker of prostatic cancer. *Glycobiology* **20**, 452–460
 14. Tateno, H., Uchiyama, N., Kuno, A., Togayachi, A., Sato, T., Narimatsu, H., and Hirabayashi, J. (2007) A novel strategy for mammalian cell surface glycome profiling using lectin microarray. *Glycobiology* **17**, 1138–1146
 15. Kaji, H., Saito, H., Yamauchi, Y., Shinkawa, T., Taoka, M., Hirabayashi, J., Kasai, K., Takahashi, N., and Isobe, T. (2003) Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat. Biotechnol.* **21**, 667–672
 16. Kuno, A., Uchiyama, N., Koseki-Kuno, S., Ebe, Y., Takashima, S., Yamada, M., and Hirabayashi, J. (2005) Evanescent-field fluorescence-assisted lectin microarray. A new strategy for glycan profiling. *Nat. Methods.* **2**, 851–856
 17. Pilobello, K. T., Slawek, D. E., and Mahal, L. K. (2007) A ratiometric lectin microarray approach to analysis of the dynamic mammalian glycome. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11534–11539
 18. Shibuya, N., Goldstein, I. J., Broekaert, W. F., Nsimba-Lubaki, M., Peeters, B., and Peumans, W. J. (1987) The elderberry (*Sambucus nigra* L.) bark lectin recognizes the Neu5Ac(alpha 2-6)Gal/GalNAc sequence. *J. Biol. Chem.* **262**, 1596–1601
 19. Kaku, H., Tanaka, Y., Tazaki, K., Minami, E., Mizuno, H., and Shibuya, N. (1996) Sialylated oligosaccharide-specific plant lectin from Japanese elderberry (*Sambucus sieboldiana*) bark tissue has a homologous structure to type II ribosome-inactivating proteins, ricin and abrin. cDNA cloning and molecular modeling study. *J. Biol. Chem.* **271**, 1480–1485
 20. Hirabayashi, J., Dutta, S. K., and Kasai, K. (1998) Novel galactose-binding proteins in Annelida. Characterization of 29-kDa tandem repeat-type lectins from the earthworm *Lumbricus terrestris*. *J. Biol. Chem.* **273**, 14450–14460
 21. Yabe, R., Suzuki, R., Kuno, A., Fujimoto, Z., Jigami, Y., and Hirabayashi, J. (2007) Tailoring a novel sialic acid-binding lectin from a ricin-B chain-like galactose-binding protein by natural evolution-mimicry. *J. Biochem.* **141**, 389–399
 22. Tateno, H., Mori, A., Uchiyama, N., Yabe, R., Iwaki, J., Shikanai, T., Angata, T., Narimatsu, H., and Hirabayashi, J. (2008) Glycoconjugate microarray based on an evanescent-field fluorescence-assisted detection principle for investigation of glycan-binding proteins. *Glycobiology* **18**, 789–798
 23. Hu, D., Kamiya, Y., Totani, K., Kamiya, D., Kawasaki, N., Yamaguchi, D., Matsuo, I., Matsumoto, N., Ito, Y., Kato, K., and Yamamoto, K. (2009) Sugar-binding activity of the MRH domain in the ER alpha-glucosidase II beta subunit is important for efficient glucose trimming. *Glycobiology* **19**, 1127–1135
 24. Tateno, H., Nakamura-Tsuruta, S., and Hirabayashi, J. (2007) Frontal affinity chromatography. Sugar-protein interactions. *Nat. Protoc.* **2**, 2529–2537
 25. Nakamura, S., Yagi, F., Totani, K., Ito, Y., and Hirabayashi, J. (2005) Comparative analysis of carbohydrate-binding properties of two tandem repeat-type Jacalin-related lectins, *Castanea crenata* agglutinin and *Cycas revoluta* leaf lectin. *FEBS J.* **272**, 2784–2799
 26. Mehmet, H., Scudder, P., Tang, P. W., Hounsell, E. F., Caterson, B., and Feizi, T. (1986) The antigenic determinants recognized by three monoclonal antibodies to keratan sulphate involve sulphated hepta- or larger oligosaccharides of the poly(N-acetylglucosamine) series. *Eur. J. Biochem.* **157**, 385–391
 27. Tabata, N., Sakuma, Y., Honda, Y., Doi, N., Takashima, H., Miyamoto-Sato, E., and Yanagawa, H. (2009) Rapid antibody selection by mRNA display on a microfluidic chip. *Nucleic Acids Res.* **37**, e64
 28. Yamashita, K., Umetsu, K., Suzuki, T., and Ohkura, T. (1992) Purification and characterization of a Neu5Ac alpha 2->6Gal beta 1->4GlcNAc and HSO3(-)->6Gal beta 1->GlcNAc specific lectin in tuberous roots of *Trichosanthes japonica*. *Biochemistry* **31**, 11647–11650
 29. Suzuki, R., Kuno, A., Hasegawa, T., Hirabayashi, J., Kasai, K. I., Momma, M., and Fujimoto, Z. (2009) Sugar-complex structures of the C-half domain of the galactose-binding lectin EW29 from the earthworm *Lumbricus terrestris*. *Acta Crystallogr. D. Biol. Crystallogr.* **65**, 49–57
 30. Hemmi, H., Kuno, A., Ito, S., Suzuki, R., Hasegawa, T., and Hirabayashi, J. (2009) NMR studies on the interaction of sugars with the C-terminal domain of an R-type lectin from the earthworm *Lumbricus terrestris*. *FEBS J.* **276**, 2095–2105
 31. Wang, Y., Yu, G., Han, Z., Yang, B., Hu, Y., Zhao, X., Wu, J., Lv, Y., and Chai, W. (2011) Specificities of *Ricinus communis* agglutinin 120 interaction with sulfated galactose. *FEBS Lett.* **585**, 3927–3934