J. Biochem. 2009;146(3)369–373
doi:10.1093/jb/mvp079

Human C21orf63 is a Heparin-binding Protein

Kanae Mitsunaga[1,*], Jun Harada-Itadani[2], Toshihide Shikanai[1], Hiroaki Tateno[3], Yuzuru Ikehara[2], Jun Hirabayashi[3], Hisashi Narimatsu[1] and Takashi Angata[1,†,‡]

[1]Glycogene Function Team; [2]Molecular Medicine Team; and [3]Lectin Application and Analysis Team, Research Center for Medical Glycoscience, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan

Received March 3, 2009; accepted May 18, 2009; published online May 26, 2009

Human C21orf63 is a type-1 transmembrane protein of hitherto unknown function, with two repeats of putative 'galactose-binding lectin domains'. By using glycan microarray analysis and other assays, we found that human C21orf63 interacts with heparin and to a lesser extent with heparan sulphate. The C-terminal galactose-binding lectin domain of C21orf63 is necessary for heparin binding. The inability of other human proteins with galactose-binding lectin domains to interact with heparin suggests that heparin binding is a unique property of C21orf63. Results of real-time polymerase chain reaction and tissue immunostaining imply that C21orf63 is expressed on epithelia of various human tissues.

Key words: C21orf63, glycan microarray, heparan sulphate, heparin, lectins.

Abbreviations: BSA, bovine serum albumin; CS, chondroitin sulphate; ELISA, enzyme-linked immunosorbent assay; GAG, glycosaminoglycan; HS, heparan sulphate; IgG, immunoglobulin G; PBS, phosphate-buffered saline; PBST, PBS containing 0.05% Tween 20; PCR, polymerase chain reaction; RT–PCR, reverse transcription PCR.

Numerous mammalian lectins have been discovered and characterized in the past (*1*). Some mammalian lectins are involved in intracellular processes such as sorting and quality control of glycoproteins. A number of lectins are found in the immune system, where they are involved in cell-to-cell interactions, recognition of foreign organisms and regulation of immune cell activity. In light of the vast structural diversity of glycan structures, however, it is likely that there are more mammalian lectins that regulate various biological processes. Better comprehension of mammalian lectin repertoires would lead to a greater understanding of glycan functions in mammals.

On the one hand, most of the known mammalian lectins belong to several protein families (e.g. the C-type lectin and galectin families) that also have members in non-mammals. On the other hand, some non-mammalian lectins and mammalian proteins share the same structural domain, though the glycan-binding capability of the latter has not been tested. These facts suggest that the mammalian proteome has been under-explored for possible glycan-binding proteins. Advancements in the molecular cloning of various non-mammalian lectins and completion of the sequencing of the human genome allow us a comprehensive bioinformatics survey of human genome for potential lectin genes. Combined with the development of glycan microarray (*2*), screening of proteins for glycan-binding properties in a high-throughput manner is now feasible.

Proteins that contain 'galactose-binding lectin domain' (PFAM accession number: PF02140) and show glycan-binding property have been found in sea urchin eggs (*3*) and in fish eggs (*4*). Through a database search, we found several human proteins containing this domain, only one of which has been tested for glycan-binding capability. In this study we expressed some of these proteins in recombinant form, tested their glycan-binding preference by glycan microarray and found that human C21orf63 protein interacts with heparin. We further characterized the glycan-recognition specificity of the protein. Our approach facilitates discovery of a lectin even when the glycan-binding specificity of the candidate is far from the expected specificity, and hence will be useful for the screening of novel lectins.

MATERIALS AND METHODS

Materials—Heparan sulphate (HS; from bovine kidney), chondroitin sulphate (CS)-A (from whale cartilage), CS-B (from pig skin), CS-C (from shark cartilage), CS-D (from shark cartilage) and CS-E (from squid cartilage) were purchased from Seikagaku Biobusiness Corporation. Heparin (from bovine lung) was purchased from Sigma.

Database Search for Potential Human Lectins—Protein domains with glycan-binding ability from all sources were searched in the PFAM database (http://www.sanger.ac.uk/Software/Pfam/) (*5*). Human genome and transcriptome data in the GenBank were surveyed for sequences containing these putative glycan-binding

*Present address: Division of Stem Cell Regulation, Center for Molecular Medicine, Jichi Medical University, 3311-1 Yakushiji, Shimotsuke, Tochigi 329-0498, Japan.
†To whom correspondence should be addressed.
Tel: +81-6-6879-8424 (ext: 2480),
E-mail: angata@sanken.osaka-u.ac.jp
‡Present address: The Institute of Scientific and Industrial Research, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan.

domains by using the sequence search algorithm that we used previously (*6*).

Cloning and Expression of Recombinant Human Proteins with Galactose-binding Lectin Domain—A full-length cDNA of C21orf63 (GenBank accession number: BC038710) was obtained via Open Biosystems. Full-length cDNAs of latrophilins 1 through 3 were cloned by reverse transcription-polymerase chain reaction (RT-PCR). cDNA segments encoding the entire extracellular domain (C21orf63) or the N-terminal part comprising the galactose-binding lectin and olfactomedin-like domains (latrophilins) of these proteins were amplified by PCR, and cloned into an AvT-EK-Fc/pcDNA3.1 vector, a derivative of EK-Fc/pcDNA3.1 (*7*). Expression vectors for C21orf63 domain-deletion mutants (ΔN and ΔC, lacking amino acids 67–160 and 164–260, respectively) and point mutants (F241A and K250A) were prepared by using a KOD Plus Mutagenesis Kit (Toyobo) in accordance with the manufacturer's instructions. Expression and purification of the recombinant proteins were carried out as described previously (*7*). A full-length C21orf63 protein expression construct (C21orf63/pcDNA3.1) was prepared by PCR amplification of the open reading frame and cloning into pcDNA3.1(–). The primers used are listed in Supplementary Table 1.

Glycan Microarray Analysis—We used version 4.2 of the glycan microarray (*2*). The layout of the array is shown in Supplementary Fig. 1. Purified recombinant proteins (40 ng) were mixed with 40 ng of Cy3-labelled goat anti-human immunoglobulin G (IgG) (Jackson ImmunoResearch) in 40 μl of assay buffer (25 mM Tris–HCl, pH 7.4, 200 mM NaCl, 1 mM $MnCl_2$, 1 mM $CaCl_2$, 1% Triton-X100), and overlaid onto the glycan microarray. After incubation for 3 h in a humidified chamber, the array wells were scanned with an evanescent microarray scanner (SC-Profiler, Moritex Corporation), and the fluorescence signal of each spot was captured and quantified.

Preparation of Glycosaminoglycan-Conjugated and Biotinylated Bovine Serum Albumin—Glycosaminoglycans (GAGs) were conjugated to bovine serum albumin (BSA) by using *N*-ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline as described previously (*8*). The coupling reagent was removed by ultrafiltration (cut off = 10 kDa), and the products were biotinylated by using Sulfo-NHS-Biotin (Pierce).

GAG-BSA-Binding Assay—The GAG-BSA-binding assay was performed as follows: each well of a 96-well plate (Nunc #269620) was coated with 100 μl of protein A (5 μg/ml; Sigma) in 50 mM sodium bicarbonate buffer, pH 9.5 at 4°C overnight, washed twice with enzyme-linked immunosorbent assay (ELISA) buffer (20 mM HEPES–NaOH buffer, pH 7.45, 125 mM NaCl, 0.02% NaN_3, 1% BSA, 1 mM $CaCl_2$, 1 mM $MgCl_2$), and blocked with ELISA buffer at room temperature for 1 h. Thereafter, the wells were sequentially incubated at room temperature with the following reagents (100 μl/well), each followed by three washes with ELISA buffer (150 μl/well): recombinant C21orf63-Fc protein or human IgG (as a negative control) in ELISA buffer (2 μg/ml) for 2 h; GAG-BSA in ELISA buffer (0.25 μg/ml) for 2 h; streptavidin-alkaline phosphatase (Jackson ImmunoResearch) in ELISA buffer (1 μg/ml) for 1 h. Finally, 100 μl of alkaline phosphatase substrate (10 mM *p*-nitrophenyl phosphate in 100 mM Na_2CO_3, 1 mM $MgCl_2$) was added to each well and colour development was monitored at 405 nm with a VersaMax microplate reader (Molecular Devices). Binding was tested in triplicate wells for each ligand.

Development of Polyclonal Antibody Against Human C21orf63—Rabbit antiserum against a segment of human C21orf63 ($C+I^{288}NFDPSGSKVLRKD^{301}$, where C+ indicates an extra cysteine residue added at the N-terminus) was prepared by a local contractor (Operon Biotechnology, Tokyo, Japan). The antibody fraction recognizing the peptide was purified by affinity purification using a peptide-conjugated agarose resin.

Staining of C21orf63-Transfected Cells with the Anti-C21orf63 Polyclonal Antibody—We transfected COS-1 cells cultured on two-well chamber slide with C21orf63/pcDNA3.1 or control plasmid (pcDNA3.1) using FuGENE 6 (Roche Applied Science). Cells were fixed 24 h after transfection with 2% formaldehyde in PBS for 10 min, blocked with 10% goat serum, 3% BSA in PBS and then incubated with the rabbit polyclonal anti-C21orf63 antibody (diluted at ~1.6 μg/ml in 0.1% BSA/PBS) at 4°C overnight. The slide was washed with PBS, incubated with an AlexaFluor-488 labelled $F(ab')_2$ fragment of goat anti-rabbit IgG (Invitrogen; 1:2,000 in PBS) for 30 min, washed with PBS and then incubated with Hoechst 33258 (Invitrogen; 1:5,000 in PBS) for 15 min for nuclear staining. The slide was washed with PBS, mounted in ProLong Gold Antifade Reagent (Invitrogen) and examined with an Axio Observer D1 microscope (Zeiss). Photomicrographs (fluorescence and differential interference contrast micrographs) were captured digitally and overlaid using AxioVision software (Zeiss) for presentation.

Immunohistochemical Analysis of Human Tissues with Anti-C21orf63 Antibody—Frozen human stomach tissues were cut at a thickness of 5 μm, mounted onto glass slides and fixed in cold acetone (–20°C) for 5 min. The slides were blocked with 5% BSA in PBS containing 0.05% Tween 20 (PBST) for 1 h and then incubated with the rabbit polyclonal anti-C21orf63 antibody (diluted at ~1 μg/ml in PBST) at 4°C for 16 h. After being washed with PBST, the slides were incubated with an AlexaFluor-488 labelled $F(ab')_2$ fragment of goat anti-rabbit IgG (Invitrogen; 1:2,000 in PBST) for 30 min, washed and then incubated with TO-PRO-3 (Invitrogen; 1:5,000 in PBST) for 10 min for nuclear staining. The slides were washed with PBST, mounted in ProLong Gold Antifade Reagent (Invitrogen) and examined with an LSM 510 confocal laser microscope (Zeiss). Photomicrographs were captured digitally and processed minimally with Photoshop (Adobe Systems) for presentation.

RESULTS

Human C21orf63 is a Heparin-Binding Protein—To test if any 'galactose-binding lectin domain' in human proteins recognizes glycans, we expressed C21orf63, latrophilins 1 and 3 in the form of Fc-fusion protein and analysed their glycan-binding ability by using

glycan microarray. We could not analyse latrophilin 2, because the recombinant protein was poorly expressed. We found none of these proteins bound to L-rhamnose or D-galactose, the two sugars that are recognized by sea urchin and salmonid fish galactose-binding lectins (*3*, *4*). However, human C21orf63, a protein with two tandem galactose-binding lectin domains (see Supplementary Fig. 2 for annotations), showed specific binding to heparin (Fig. 1). Neither latrophilin 1 nor latrophilin 3 showed significant binding to any of the glycans, including heparin, in analyses using the glycan microarray or the ELISA-like assay described below (data not shown).

Human C21orf63 Shows Specific Interaction with Heparin and not with Chondroitin Sulphate—To confirm the binding specificity of C21orf63, we conjugated BSA to various GAGs, and used the resultant probes (GAG-BSA) in an ELISA-like binding assay with recombinant C21orf63 protein. Heparin-BSA showed strong binding to C21orf63, whereas HS-BSA showed much weaker binding (Fig. 2A). In contrast, none of the CS-BSA probes showed detectable binding. These results essentially confirmed the results obtained by glycan microarray analysis.

We further analysed the GAG-binding specificity of C21orf63 by a competitive inhibition assay, in which various GAGs were added to the assay system as competitors. As shown in Fig. 2B, heparin almost completely blocked heparin-BSA binding to C21orf63 at the lowest concentration used (3 μg/ml), whereas most other GAGs showed only moderate inhibition at the highest concentration tested (20 μg/ml). This result confirms that C21orf63 shows a strong preference for heparin.

We observed no significant change of ligand binding when we omitted divalent cations from the assay buffer (data not shown), and concluded that glycan binding by C21orf63 protein is not dependent on divalent cations. This conclusion is consistent with the previous report on fish lectins belonging to this family (*4*).

The C-terminal Domain of C21orf63 is Required for Heparin Binding—To determine which part of C21orf63 is involved in its interaction with heparin, we prepared mutant C21orf63 proteins with either the N- or the C-terminal galactose-binding lectin domain deleted. In our ELISA-like binding assay, the deletion mutant

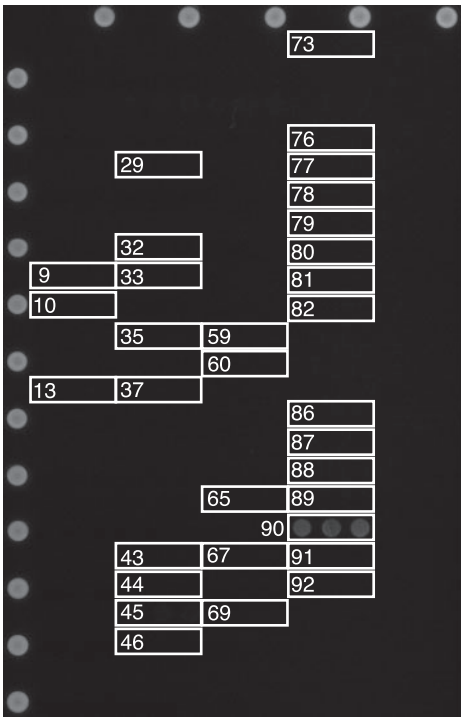

Fig. 1. **Glycan microarray analysis of C21orf63–glycan interactions.** Recombinant C21orf63-Fc protein was applied to a glycan microarray, and binding was detected with Cy3-labelled anti-human IgG antibody. A distinct binding signal was observed only for heparin (slot 90). C21orf63 did not show any binding to L-rhamnose (slot 92), glycans with terminal β-D-galactose (slots 10, 13, 29, 32, 33, 35, 37, 43-46, 59, 60, 67, 69, 73) or terminal α-D-galactose residues (slots 9, 65, 76–82), or glycosaminoglycans other than heparin (slots 86–89, 91). A full description of the glycans spotted on the array is given in Supplementary Figure 1.

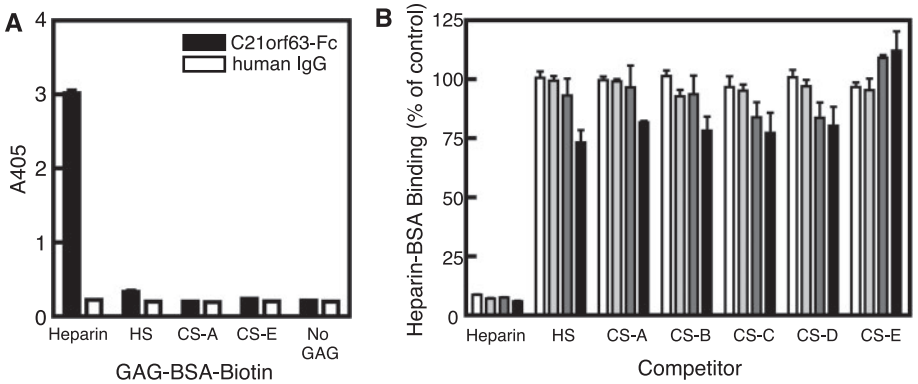

Fig. 2. **Binding of C21orf63 to glycosaminoglycans.** (A) Binding of glycosaminoglycan-conjugated and biotinylated BSA (GAG-BSA) to recombinant C21orf63-Fc protein was quantified by an ELISA-like method. Mean colorimetric readouts at 405 nm (A405) are plotted. Human IgG was used as a control for C21orf63-Fc, and biotinylated BSA (without a glycosaminoglycan chain attached) was used as a control for GAG-BSA. (B) Binding of heparin-BSA to immobilized C21orf63-Fc was measured in the presence of free glycosaminoglycans (3, 5, 10 or 20 μg/ml, represented by open, light grey, dark grey and filled bars, respectively) as competitors, and is plotted as the binding signal intensity relative to that observed in the absence of competitor. Error bars represent standard deviations.

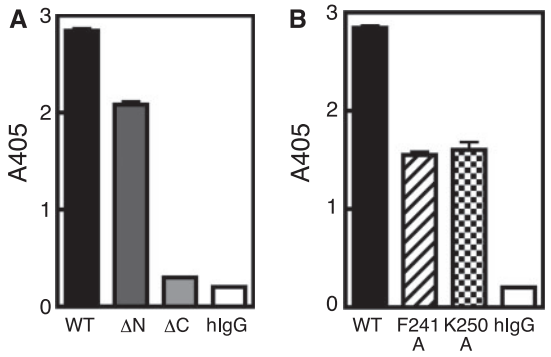

Fig. 3. **Structural elements of C21orf63 involved in heparin binding.** Heparin-BSA binding to mutant C21orf63-Fc was analysed by the ELISA-like assay. (A) Heparin-BSA binding to wild type C21orf63 (filled bar), the deletion mutant lacking the N-terminal domain (ΔN, dark-grey bar) and the one lacking the C-terminal domain (ΔC, light-grey bar). (B) Heparin-BSA binding to wild type C21orf63 (filled bar) and point mutants at conserved amino-acid residues in the C-terminal domain (F241A or K250A, represented by hatched and checkered bars, respectively). Human IgG (white bar) was used as a control. The mean colorimetric readouts at 405 nm (A405) are plotted. Error bars represent standard deviations.

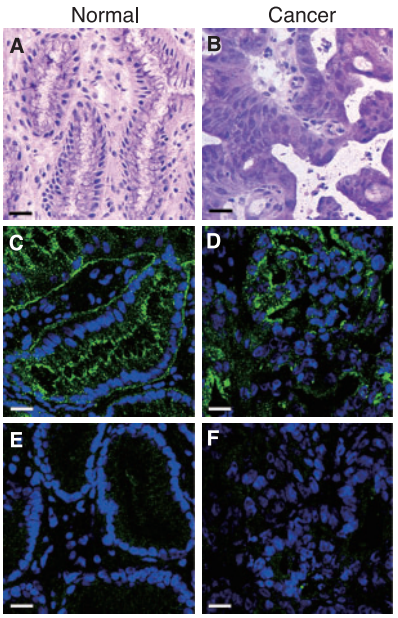

Fig. 4. **C21orf63 protein localization in human stomach mucosa.** (A, C and E) are photomicrographs of normal gastric mucosa; (B, D and F) are those of gastric cancer mucosa. Tissue sections were stained by standard haematoxylin-eosin staining (A and B), with anti-C21orf63 antibody (C and D) or with control rabbit IgG (E and F). Green fluorescence signals in (C) and (D) represent antibody binding to C21orf63 protein; blue fluorescence signals represent nuclei. Scale bars: 20 μm.

lacking the C-terminal domain (ΔC) showed almost a complete loss of binding to heparin-BSA, whereas the one lacking the N-terminal domain (ΔN) showed only a moderate reduction in heparin-BSA binding as compared with the full-length protein (Fig. 3A).

To analyse the amino-acid residues involved in heparin binding, we prepared two mutant C21orf63 proteins (F241A and K250A) that had point mutations in the C-terminal domain at amino-acid residues that are conserved among various galactose-binding lectin domains (Supplementary Fig. 3). Both proteins showed a partial reduction in heparin binding (Fig. 3B). Taken together, these results imply that the C-terminal galactose-binding lectin domain is involved in heparin binding, and that amino-acid residues highly conserved among family members are not absolutely required.

C21orf63 is Expressed on Epithelial Cells of Stomach—To analyse the expression pattern of C21orf63, we performed a quantitative PCR analysis of RNA samples from various human tissues. We found that C21orf63 mRNA is ubiquitously expressed in various tissues (Supplementary Fig. 4). We then prepared a polyclonal antibody against C21orf63, confirmed its specificity by staining of the cells transfected with full-length C21orf63 expression construct or with the empty vector (Supplementary Fig. 5), and used it to probe frozen human stomach tissue sections. We found that C21orf63 protein is expressed on epithelial cells of gastric mucosa (Fig. 4C), and that its expression is maintained after malignant transformation (Fig. 4D). Taken together, these results suggest that C21orf63 may be widely expressed on the epithelial cells of human tissues.

DISCUSSION

We have described in this article the heparin-binding property of C21orf63. Although we expected that some of the human proteins with galactose-binding lectin domains would recognize glycan, we detected only the heparin binding of C21orf63. This observation can be interpreted in two different ways. One interpretation is that the affinities between the proteins and sugars might not have been high enough for detection by the method we used. In fact, the galactose-binding lectin domain of mouse latrophilin 1 was recently shown to interact with L-rhamnose at low affinity ($K_d = 1.8$ mM) by using nuclear magnetic resonance, whereas the analyses using a glycan microarray failed to detect its binding to any glycan including L-rhamnose (*9*). The failure to detect weak protein–carbohydrate interactions may be an inherent limitation of current glycan micorarray technology. An alternative interpretation is that the structural varieties of the ligands spotted on the glycan microarrays are not sufficiently broad. Nevertheless, our approach made it possible to discover the glycan-binding property of C21orf63 even though its glycan-binding specificity is different from what would be predicted from its sequence. Our approach, combined with the future expansion of ligands on the array, will be useful for the discovery of novel lectins.

We found that heparin binding is a unique property of C21orf63 that is not shared with the other human proteins with galactose-binding lectin domains we have tested (i.e. latrophilins 1 and 3). The observation that the C-terminal galactose-binding lectin domain is required for heparin binding, whereas the highly conserved residues of the galactose-binding lectin domain of C21orf63 are not absolutely required, implies that the interaction is mediated by other structural element(s) in the

C-terminal domain. In this regard, we noticed that the C-terminal domain of C21orf63 contains many positively charged amino-acid residues (14 out of 91 amino acids are lysine or arginine residues, whereas the galactose-binding lectin domains of human latrophilins contain seven to eight lysine or arginine residues), which may contribute to the interaction of this protein with heparin.

EVA-1, the putative C21orf63 ortholog in *Caenorhabditis elegans*, forms a complex with SAX-3 (*C. elegans* counterpart of *Drosophila* and mammalian Robo proteins) and functions as a co-receptor for SLT-1 (*C. elegans* counterpart of *Drosophila* and mammalian Slit proteins) (*10*). The interactions between EVA-1 and SLT-1 or SAX-3 are reported to be independent of glycan recognition by EVA-1, based on the facts that some sugars (galactose, glucose or lactose) do not inhibit EVA-1 binding to SLT-1 or SAX-3. However, interactions with other oligosaccharides or GAGs have not been tested. Meanwhile, it is known that Robo and Slit proteins form ternary complex with HS (*11*, *12*). It is tempting to speculate that C21orf63 forms a complex with heparin/HS and Robo and serves as a Slit receptor.

SUPPLEMENTARY DATA

Supplementary data are available at *JB* online.

ACKNOWLEDGEMENT

We thank Mr Kazunori Nakamura and Mr Atsushi Mori for technical assistance.

FUNDING

'Support of Young Researchers with a Term' from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to T.A.); 'Development of Novel Diagnostic and Medical Applications through Elucidation of Sugar Chain Functions' from the New Energy and Industrial Technology Development Organization (to Y.I., J.H. and H.N.).

CONFLICT OF INTEREST

None declared.

REFERENCES

1. Sharon, N. and Lis, H. (2003) *Lectins*, Kluwer Academic Publishers, Dordrecht/Boston
2. 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
3. Ozeki, Y., Matsui, T., Suzuki, M., and Titani, K. (1991) Amino-acid sequence and molecular characterization of a D-galactoside-specific lectin purified from sea urchin (*Anthocidaris crassispina*) eggs. *Biochemistry* **30**, 2391–2394
4. Tateno, H., Saneyoshi, A., Ogawa, T., Muramoto, K., Kamiya, H., and Saneyoshi, M. (1998) Isolation and characterization of rhamnose-binding lectins from eggs of steelhead trout (*Oncorhynchus mykiss*) homologous to low density lipoprotein receptor superfamily. *J Biol. Chem.* **273**, 19190–19197
5. Finn, R.D., Tate, J., Mistry, J., Coggill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L., and Bateman, A. (2008) The Pfam protein families database. *Nucleic Acids Res.* **36**, D281–D288
6. Kikuchi, N. and Narimatsu, H. (2006) Bioinformatics for comprehensive finding and analysis of glycosyltransferases. *Biochim. Biophys. Acta* **1760**, 578–583
7. Angata, T., Tabuchi, Y., Nakamura, K., and Nakamura, M. (2007) Siglec-15: an immune system Siglec conserved throughout vertebrate evolution. *Glycobiology* **17**, 838–846
8. Nishioka, S., Aikawa, J., Ida, M., Matsumoto, I., Street, M., Tsujimoto, M., and Kojima-Aikawa, K. (2007) Ligand-binding activity and expression profile of annexins in *Caenorhabditis elegans*. *J Biochem.* **141**, 47–55
9. Vakonakis, I., Langenhan, T., Promel, S., Russ, A., and Campbell, I.D. (2008) Solution structure and sugar-binding mechanism of mouse latrophilin-1 RBL: a 7TM receptor-attached lectin-like domain. *Structure* **16**, 944–953
10. Fujisawa, K., Wrana, J.L., and Culotti, J.G. (2007) The slit receptor EVA-1 coactivates a SAX-3/Robo mediated guidance signal in *C. elegans*. *Science* **317**, 1934–1938
11. Hohenester, E., Hussain, S., and Howitt, J.A. (2006) Interaction of the guidance molecule Slit with cellular receptors. *Biochem. Soc. Trans.* **34**, 418–421
12. Hussain, S.A., Piper, M., Fukuhara, N., Strochlic, L., Cho, G., Howitt, J.A., Ahmed, Y., Powell, A. K., Turnbull, J.E., Holt, C.E., and Hohenester, E. (2006) A molecular mechanism for the heparan sulfate dependence of slit-robo signaling. *J Biol. Chem.* **281**, 39693–39698