J. Biochem. **141,** 389–399 (2007)
doi:10.1093/jb/mvm043

Tailoring a Novel Sialic Acid-Binding Lectin from a Ricin-B Chain-like Galactose-Binding Protein by Natural Evolution-Mimicry

Rikio Yabe[1,2], Ryuichiro Suzuki[1], Atsushi Kuno[1,*], Zui Fujimoto[3], Yoshifumi Jigami[1,2] and Jun Hirabayashi[1]

[1]Research Center for Glycoscience, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki 305-8568, Japan; [2]Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8572, Japan; and [3]Department of Biochemistry, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan

Received December 2, 2006; accepted January 6, 2007; published online January 18, 2007

Sialic acid (Sia) is a typical terminal sugar, which modifies various types of glycoconjugates commonly found in higher animals. Its regulatory roles in diverse biological phenomena are frequently triggered by interaction with Sia-binding lectins. When using natural Sia-binding lectins as probes, however, there have been practical problems concerning their repertoire and availability. Here, we show a rational creation of a Sia-binding lectin based on the strategy 'natural evolution-mimicry', where Sia-binding lectins are engineered by error-prone PCR from a Gal-binding lectin used as a scaffold protein. After selection with fetuin–agarose using a recently reinforced ribosome display system, one of the evolved mutants SRC showed substantial affinity for α2-6Sia, which the parental Gal-binding lectin EW29Ch lacked. SRC was found to have additional practical advantages in productivity and in preservation of affinity for Gal. Thus, the developed novel Sia-recognition protein will contribute as useful tools to sialoglycomics.

Key words: glycomics, natural evolution, ribosome display, ricin-B chain, Sia-binding lectin.

Abbreviations: EW29, earthworm 29-kDa lectin; EW29Ch, C-terminal domain of EW29; EW29Nh, N-terminal domain of EW29; FAC, frontal affinity chromatography; PA, pyridylaminated; pNP, *p*-nitrophenyl; SRC, Sia-recognition EW29Ch.

Complex glycans are synthesized in various biological contexts, especially in those of higher animals including humans. Sialic acids (Sia) represent a group of non-reducing terminal sugars in mammalian glycoconjugates, possessing great structural diversity and biological significance. Variation in Sia derivatives, *e.g.* in linkage and positional isomers, greatly increases the range of glycan structural complexity and functional diversity. The latter includes regulatory roles in cell–cell interaction, cell-signalling and viral infections (*1*). Striking changes in sialylation features on tumour cells have been shown to serve as useful 'tumour markers' (*2*). In inflammation, the first step of leucocytic transmigration into endothelial cells is a cell-adhesion event mediated by the specific recognition of sialyl Lewis X by selections (*3*). In differentiation and development, SSEA-4 (Siaα2-3Galβ1-3GalNAcβ1-3Galα1-4R) is a well-known biomarker (*4*). Recently, structural glycomics using mass spectrometry (MS) has been used to elucidate functional aspects of glycans (*5*). However, direct analysis of sialoglycans by MS is technically difficult, because of the low ionization efficiency of this 9-carbon acidic sugar, and substantial decomposition of Sia linkages.

Lectin-based glycotechnology is expected to be an alternative means to profile complex features of glycoproteins in a rapid and comprehensive manner (*6–8*). In this regard, various types of lectin microarray have now become available for analysing glycan profiles expressed on individual glycoproteins and even cell and tissue extracts (*9–15*). However, these systems are incomplete in terms of the repertoire of Sia-binding lectins: this is partly because stable production in bacteria of mammalian (and thus, endogenous) Sia-binding lectins is difficult. In addition, few Sia-binding lectins of plant origin are available. Strong support for a biotechnological approach appears to be necessary in order to develop a satisfactory repertoire of Sia-binding lectins.

In this study, we describe a novel lectin generated by a rational strategy mimicking the natural process of lectin evolution (Fig. 1). In essence, most lectins in nature are thought to have developed from a more primitive molecule in the order Glc/Man- to Gal-specific lectins. Applying this hypothesis, Sia-specific lectins should have emerged in the last step of evolution to recognize 'Sia-modified Gal' residue (*16*). In fact, certain trace features of Gal-recognition remain in the naturally occurring Sia-binding lectins, *e.g. Sambucus nigra* agglutinin (SNA), *Sambucus sieboldiana* agglutinin (SSA), *Polyporus squamosus* lectin (PSL), *Maackia amurensis* leukoagglutinin (MAL) and galectins (*17–24*). Therefore, valuable Sia-recognition

*To whom correspondence should be addressed. Tel: +81-29-861-3187, Fax: +81-29-861-3125, E-mail: atsu-kuno@aist.go.jp

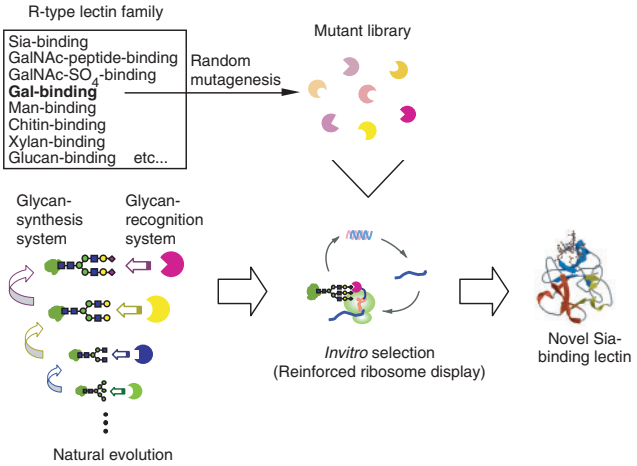

Fig. 1. **Schematic diagram of the present strategy 'natural evolution-mimicry'.** The basic concept of the methodology is to create functional Sia-binding lectins starting from a naturally occurring Gal-binding lectin as a scaffold protein. The actual procedure consists of random mutagenesis by error-prone PCR and *in vitro* selection by the reinforced ribosome display method. In this study, a Gal-binding lectin, designated EW29Ch belonging to the R-type lectin family was used as a scaffold protein.

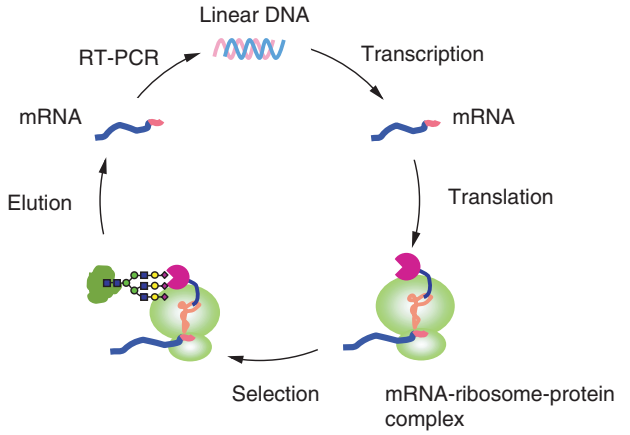

Fig. 2. **Schematic representation of the reinforced ribosome display system.** After linear DNA is transcribed to mRNA, it is subjected to translation for a short period. The translation is terminated by settling the reaction tube on ice and a Mg^{2+}-containing solution is added. The resulting mRNA–ribosome–protein complex is used for selection with fetuin–agarose. The mRNA coupled with both ribosome and synthesized protein is eluted with an EDTA-containing solution. The resulting dissociated mRNA is reverse-transcribed, and the resulting cDNA is amplified by PCR.

lectins should be easily acquired from the ancestral Gal-binding lectins by molecular evolution methodology.

This nature-mimicking strategy has been developed into a high-throughput technology by combination of error-prone PCR with a recently reinforced ribosome display system [Fig. 2; originally developed by Hanes *et al.* (*25*)]. From a practical viewpoint, the choice of an initial scaffold protein is crucial to the success of the strategy, as it must be sufficiently stable and productive in a bacterial expression system. In this study, we chose earthworm 29-kDa Gal-binding lectin (EW29) to fulfill the requirements (*26*). After enrichment with fetuin (a representative serum sialoglycoprotein) as a selection probe, one of the derived fetuin-binding lectins (designated SRC) was found to show substantial affinity for α2,6-sialylated glycans, which its parental molecule completely lacked. Notably, SRC retained lowered but detectable affinity for Gal, reminiscent of molecular evolution of Gal-binding lectins. X-ray crystallographic analysis of SRC strongly supported our hypothesis: *i.e.* modification of the original lactose-binding pocket resulted in the formation of an extended sub-recognition site for α2-6Sia. As far as we know, this is the first example of the application of molecular evolution to produce novel Sia-binding lectins. These findings have both academic and practical implications for future approaches to molecular design of carbohydrate-recognition molecules.

MATERIALS AND METHODS

Construction of the Plasmid, pRARE/ew29ch—For *in vitro* selection, Gene III gene containing an additional region utilized as a spacer (0.4 bp) was inserted into pET-27b vector (Qiagen) via *Xho*I and *Nhe*I sites (designated pRARE vector). EW29Ch gene was ligated into pRARE vector via *Nde*I and *Xho*I sites. His-*So*CBM13β′ mutant gene from *Streptomyces* xylanase was inserted into pRARE vector via *Xba*I and *Xho*I sites (pRARE/*his-socbm13*β′) (*27*).

Preparation of a DNA library—Open reading frame of EW29Ch into pRARE/*ew29ch* was randomly mutated by error-prone PCR with 100 ng/ml pRARE/*ew29ch*, Master Mix (Promega), 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 1 μM T7P-for primer (5′-TAATACGACTCACTATAGGGG AATTGTG-3′) and 1 μM gene3-rev primer (5′-CCACCTCC CTCACCGCCGGTAGAAGATATC-3′) (2 min at 98°C, followed by 30 cycles of 30 s at 98°C, 30 s at 65°C and 45 s at 72°C, with a final 5 min at 72°C) (*28*). A spacer fragment was amplified by PCR with 200 ng/ml of pRARE/*his-socbm13*β′, 1 μM of the forward primer, 'gene3-for' (5′-CTCGAGGGGATATCTTCTACCGGCGGA GGA-3′), and 1 μM of the reverse primer, 'RARE-rev' (5′-GCTAGCAACCCGAACCCGTCCCCGTGAATT-3′).
The synthesized DNA fragments were treated with *Dpn*I and then were blunt-ended by T4 DNA polymerase. Using an overlap PCR method (*29*), the library fragment was fused to the spacer fragment. The overlap PCR product was thoroughly amplified using T7P-for primer and RARE-rev primer (5′-GCTAGCAACCCGAACCCGT CCCCGTGAATT-3′). The synthesized fragments were purified with QIAquick PCR purification Kit (Qiagen) and were transcribed with RiboMAX™ Express Large Scale RNA production Systems-T7 (Promega). The resulting mRNA was purified with RNeasy (Qiagen) and was used immediately for selection.

Selection with Fetuin–agarose—*In vitro* translation was performed with the mRNA library using *Escherichia coli* S30 Extract System for Linear Templates (Promega) for 10 min at 37°C. The translation reaction was stopped by placing on ice for 5 min and by diluting with ice-cold binding buffer (200 mM sodium

phosphate buffer, 150 mM NaCl, 50 mM $MgCl_2$ and 1% Tween20, pH 5.7). Immediately after centrifugation at 10,000 × *g* for 5 min at 4°C, the supernatant was added to the pre-mixed solution containing 0.2 mg/ml fetuin-immobilized agarose, 1% BSA, 4 U/μl RNasin (Promega) in the binding buffer, and the reaction tube was rotated at 4°C for 10 min. After centrifugation, the slurry was washed with the binding buffer. The mRNA coupled with both ribosome and synthesized protein was dissociated with elution buffer (200 mM sodium phosphate buffer, 150 mM NaCl and 50 mM EDTA, pH 5.7) and the resulting mRNA fragment was purified with RNeasy. The purified RNA was reverse-transcribed and amplified with a one-step RNA PCR kit (AMV) (TaKaRa) using the 'T7P-for' and 'RARE-rev' primer set. The derived DNA fragment was used in the next selection. After the fourth selection cycle, the RT-PCR product was re-amplified by PCR using the 'T7P-for' and 'gene3-rev' primer set, and the resulting DNA fragments were inserted into pET-27b vector (Novagen) via *Nde*I and *Xho*I sites and were subjected to sequence analysis.

Protein Expression and Purification—Expression and purification of selected mutants were performed essentially as described previously (*30*). The selected mutants were expressed into *E. coli* BL21-CodonPlus (DE3)-RIL strain (Stratagene), and were purified by lactose–agarose (2 ml column volume) (Sigma). The purified mutant proteins were desalted and concentrated with an Amicon Ultra–15 (5000 MWCO, Millipore).

Lectin Microarray Analysis—Lectin microarray was performed as described previously (*9*). Cy3-labeled asialofetuin was prepared from Cy3-labeled fetuin by Sialidase A (Prozyme) digestion. Fifty μl of Cy3-labeled glycoprotein solution in the probing buffer (1% Triton X, 1 mM Ca^{2+}, 1 mM Mn^{2+} and 0.5 M glycine in TBS) was applied to the lectin array on the glass slide, which was incubated at 20°C for 8 h. The reaction solution was discarded and the glass slide was rinsed twice with the probing buffer. A fluorescence image of the microarray was acquired by an evanescent-field fluorescence scanner, GTMASScan III (Nippon Laser & Electronics lab. Nagoya, Japan).

Frontal Affinity Chromatography (FAC)—FAC was performed using an automated machine, FAC-1 (Shimadzu) as described previously (*31*). Flow rate and column temperature were kept at 0.125 ml/min and 25°C, respectively. Elution of pyridylamine (PA)- and 2-aminobenzoic acid (AA)-labelled glycans and fetuin was monitored with a fluorescence detector (excitation/emission wave lengths: 310/380, 335/410 and 285/350 nm for PA, AA and fetuin, respectively). Elution of pNP-glycans was monitored with a UV detector (280 nm). Dissociation constants (K_d) and effective ligand contents (B_t) in FAC were determined as described previously (*31*, *32*). FAC competition assay was performed using α2,6-sialylated bi-antennary *N*-glycan (**503**, see Fig. 3) and fetuin in the presence of 10–200 mM lactose (Wako).

Crystallization—pET27/SRCstop vector, in which a stop codon was inserted in the 3′ terminus of the SRC gene was constructed. Ligand-free SRC, SRC/lactose and SRC/6′-sialyllactose complexes were crystallized by mixing 5 μl of 15 mg/ml protein solution with/without 15 mg/ml lactose and 6′-sialyllactose (Funakoshi), and an equal volume of precipitant solution (conditions for ligand-free SRC and SRC/lactose complex; 0.2 M ammonium sulphate, 0.1 M sodium acetate trihydrate pH4.6, 30% PEG MME 2000; those for SRC/6′-sialyllactose complex; 0.2 M NaCl, 0.1 M phosphate–citrate pH4.2, 20% PEG 8000) equilibrated against 500 μl of the reservoir solution by the sitting- or hanging-drop vapour-diffusion methods. Crystals of SRC and two complexes were obtained at 4°C within 5 days.

Data Collection—X-ray diffraction experiments were conducted using a R-AXIS VII imaging-plate X-ray detector with CuK_α X-ray diffractometer (Rigaku). Data collections were processed and scaled using DENZO and SCALEPACK from HKL2000 package (*33*). All crystals belonged to monoclinic space group $P2_1$ and diffracted beyond ~2.0 Å. Data collection and refinement statistics are summarized in Table 1. Matthews coefficient V_M values for these crystals were calculated to be 2.2 $Å^3$/Da and these crystals corresponded to a solvent content of ~47.1, 47.0 and 45.5%, respectively (*34*). An initial model of ligand-free SRC was built by the molecular replacement method using MolRep in CCP4 suite with structure of EW29Ch as a search model (PDB entry 2D12) (*35*, *36*). All crystallographic refinements were performed using CNS and the parameter set of Engh and Huber (*37*, *38*). Manual model rebuilding was conducted with XtalView (*39*). Model building of SRC/sugar complexes was conducted as described earlier with the structure of SRC (PDB entry 2DRY) as a search model. The quality of all models was analysed with PROCHECK (*40*). The coordinates for both free and complex forms of SRC have been deposited in the RCSB Protein Data Bank (accession numbers 2DRY, 2DRZ and 2DS0).

RESULTS

Selection for Sia-Binding Lectins by Natural Evolution-Mimicry—To create Sia-binding lectins using the proposed strategy—natural evolution-mimicry, EW29 was chosen as a scaffold protein (*26*). EW29, is a Gal-binding lectin belonging to the ricin-B chain (R-type lectin) family. The molecule comprises two homologous domains, *i.e.* N-terminal (N-half, EW29Nh) and C-terminal domains (C-half, EW29Ch), each divided into three subdomains designated α, β and γ. Recent X-ray crystallographic analysis of EW29Ch in complex with lactose demonstrated that it had two Gal-binding pockets in subdomains α and γ. On the other hand, EW29Nh has a much weaker sugar-binding affinity, though it is required for high-affinity binding. From a practical viewpoint, EW29Ch has a considerable advantage over other proteins in terms of productivity (>10 mg soluble protein is obtained from a 1 l culture). In addition, R-type lectins have a wide range of sugar-binding specificities. Therefore, EW29Ch is expected to be well tailored for the development of Sia-binding lectins.

As a selection probe, we chose fetuin having α2,6-sialylated tri-antennary *N*-glycan as a major

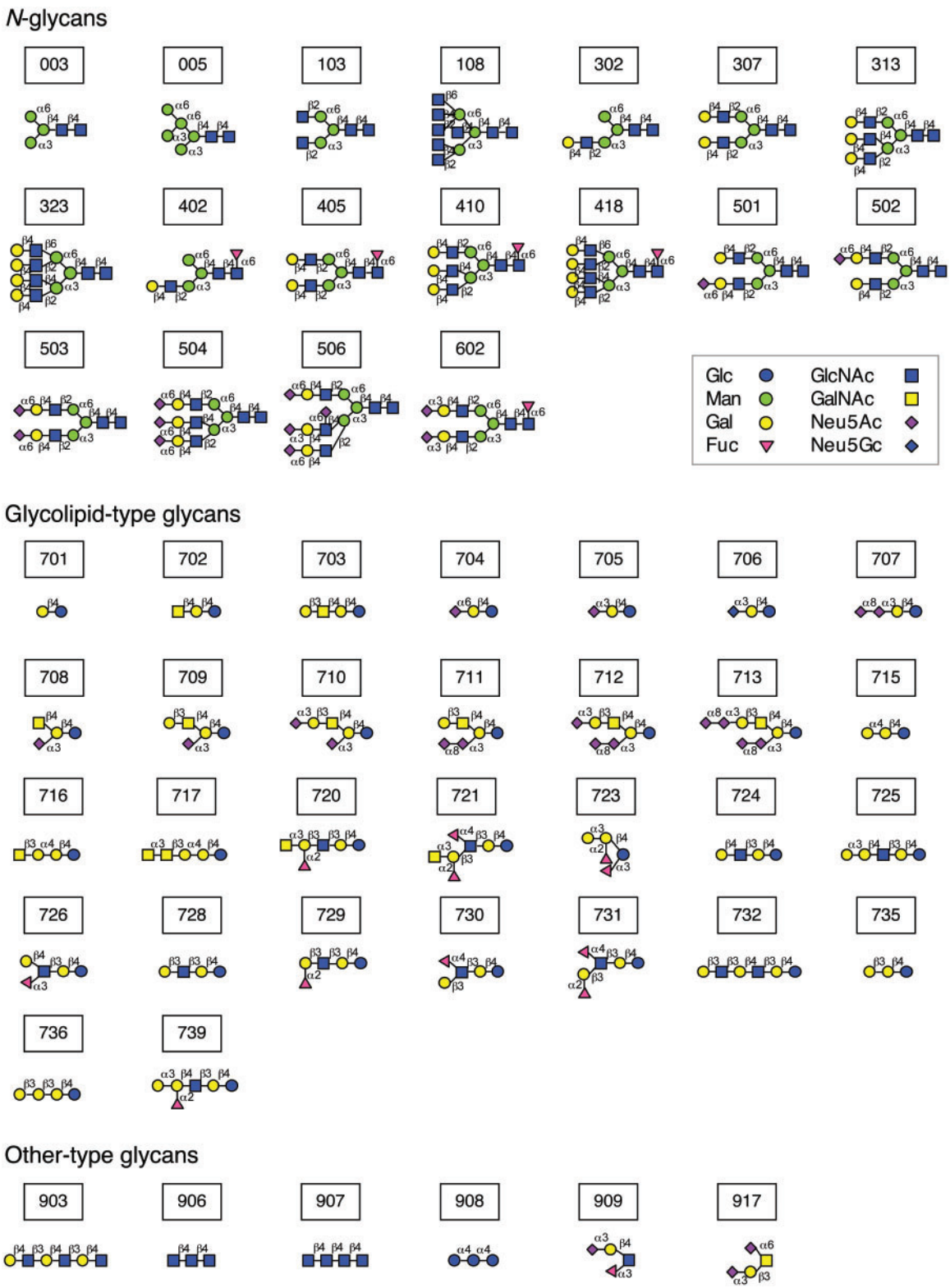

Fig. 3. **Schematic representation of 54 oligosaccharide structures used for FAC in this study.** Symbols for monosaccharides (CFG format) are shown in the box.

Table 1. **Data collection and refinement statistics of ligand-free SRC and SRC in complexes with lactose and 6′-sialyllactose.**

	Ligand-free	Lactose	6′-sialyllactose
Data collection			
Space group	$P2_1$	$P2_1$	$P2_1$
Cell dimensions			
a (Å)	40.2	40.2	40.4
b (Å)	89.6	89.3	87.7
c (Å)	41.7	41.6	41.4
α (°)	90.0	90.0	90.0
β (°)	116.4	116.3	117.4
γ (°)	90.0	90.0	90.0
Wavelength (Å)	1.5418	1.5418	1.5418
Temperature (K)	100	100	100
Resolution (Å)	50–1.80 (1.86–1.80)	50–2.20 (2.28–2.20)	50–1.80 (1.86–1.80)
R-merge (%)	3.7 (17.8)	11.1 (27.3)	3.6 (9.2)
Average $I/\sigma(I)$	70.0 (18.4)	60.6 (25.3)	76.8 (34.1)
Completeness (%)	87.1 (82.3)	90.7 (87.6)	96.2 (93.4)
Redundancy	6.5 (6.6)	6.6 (6.8)	6.9 (6.6)
Refinement			
Resolution (Å)	22.6–1.80	27.4–2.20	22.9–1.80
No. reflections	21367	12244	22852
R_{work}/R_{free} (%)	23.4/26.3	21.5/26.9	20.9/24.1
No. of atoms			
Protein	2040	2040	2040
Ligand/ion	40	66	111
Water	252	166	211
B-factors (Å^2)	25.1	31.4	17.4
R.M.S. deviations			
Bond length (Å)	0.005	0.006	0.005
Bond angles (°)	1.2	1.3	1.3
Ramachandran plot (%)			
Most favoured region	85.0	81.8	87.3
Additional allowed region	15.0	18.2	12.7
Generously allowed region	0.0	0.0	0.0
Disallowed region	0.0	0.0	0.0

Note: Values in parentheses refer to the highest resolution shell.

component, as well as short fragments of *O*-glycans known as sialyl Tn and T antigens. For selection, fetuin was immobilized to agarose at a concentration of 10 mg/ml. For construction of a library of randomly point-mutated EW29Ch genes, error-prone PCR was performed under the optimized conditions to attain both high error-incorporation and substantial amplification: *i.e.* 5 mM Mg^{2+}, 0.5 mM Mn^{2+} and low-fidelity DNA polymerase were used for amplification of 1 μg of the DNA fragment (*28*). Theoretical diversity and average mutation rate were estimated to be $>10^{12}$ and 7.5 bps/gene, respectively.

After enrichment cycles with fetuin–agarose (Fig. 2), the resultant libraries were subjected to PCR under the normal conditions using a specific primer set (T7P-for and RARE-rev), which encompassed an entire region of the EW29Ch gene (~1 kbp). The DNA fragments thus amplified were found to be substantially enriched as selection cycles proceeded (Fig. 4), when compared with a control EW29Ch. After the fourth cycle, the amplified fragments were inserted into an expression vector (pET-27b) via restriction sites (*Nde*I/*Xho*I), and the

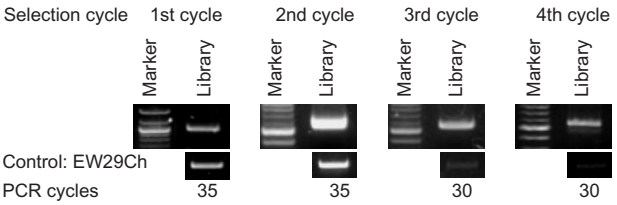

Fig. 4. **Check of RT-PCR products by agarose gel electrophoresis.** After each selection cycle with fetuin–agarose, an equal portion of RT-PCR products was subjected to analysis. A target fragment became substantially enriched as selection cycles proceeded. Note that PCR cycles required for target amplification were reduced in the third and the fourth cycles. EW29Ch was used as a control.

derived clones were extensively characterized. Sequencing of 40 such clones revealed considerable diversity at both nucleotide (8.8 mutations/gene) and amino acid (5.8) levels (Fig. 5). No overlap clone was found between the third (*20*) and the fourth (*40*) cycles (data not shown), while two (4F20 and 4F61, designated SRC later) of the latter were identical and had six amino acid changes; *i.e.* E148G, I227N, D230G, I231V, E237G and G239S (Fig. 6A). Of the 40 mutants, 31 were successfully expressed in *E. coli* as a major component of soluble extract, and 10 of them were bound to a lactose–agarose column and eluted with 0.2 M lactose. The selected 10 mutant lectins were then analysed by FAC using 12 representative glycans (**003**, **103**, **105**, **307**, **313**, **323**, **503**, **504**, **602**, **724**, **728** and **736** in Fig. 3) (*31*). The above clones, 4F20 and 4F61, showed substantial affinity for α2,6-sialylated glycans, while the others and the parental EW29Ch did not (data not shown). We focused further analysis on this α2,6Sia-binding EW29Ch mutant (designated SRC, Sia-recognition EW29Ch).

Sugar-binding Specificity of SRC—As described earlier, SRC was successfully expressed in *E. coli* as a soluble form (>20 mg/l culture), and was purified on lactose–agarose. To investigate its binding features, we performed lectin microarray analysis (*9*). SRC, as well as SNA (α2,6Sia-binding lectin) belonging to the R-type lectin family, were spotted on the same surface-activated glass slide at appropriate concentrations (2 and 1 mg/ml, respectively). As probes, Cy3-labeled fetuin, sialidase-treated fetuin (asialofetuin) and bovine serum albumin (BSA, negative control) were used at a concentration of 1 μg/ml. Intense signals were observed on SRC when Cy3-fetuin (Sia-terminated) was used as a probe, whereas less intense signals were observed for Cy3-asialofetuin (Gal-terminated; Fig. 6B). On the other hand, EW29Ch showed specific binding for asialofetuin, but not at all for fetuin. This indicated that SRC had acquired substantial affinity for Sia, which the parental EW29Ch completely lacked. Notably, SRC retains weakened but detectable affinity for Gal. The results for SNA were consistent with the previous reports (*17*, *18*).

To characterize the Sia-binding properties of SRC in greater detail, we further carried out FAC analysis using 54 glycans (Fig. 3) as described previously (*31*). The effective ligand content (B_t) of the SRC-column used,

and the dissociation constant (K_d) for pNP-β-6′-sialyllactose were determined to be 1.1 nmol and 57 μM, respectively (Table 2 and Supplementary Figure 1). Overall sugar-binding features of SRC and EW29Ch are summarized in Fig. 7A and B, respectively, in terms of affinity constant (K_a) ($K_a = 1/K_d$; for K_d values, see Supplementary Table 1). SRC showed high affinity for a series of α2,6-sialylated *N*-glycans (**501**–**506**). The binding affinity increased in the order of the branching number; *i.e*. α2,6-sialylated mono- (**501**, $K_d = 69$ μM; **502**, 70 μM), bi- (**503**, 40 μM) and tri-antennary (**504**, 27 μM) *N*-glycans. As observed in lectin microarray experiments, SRC showed weakened but detectable affinity for *N*-glycans having exposed Gal units (**302**–**418**; 99–400 μM). Affinity for α2,6-sialylated bi-antennary glycan (**503**, 40 μM) was 7.2-times stronger than its galactosylated form (**307**, 286 μM). On the other hand, parental EW29Ch showed no affinity for α2,6-sialylated *N*-glycans (**503**–**506**), whereas it showed high affinity for a series of galactosylated *N*-glycans (**302**–**418**; 1.9–13 μM). Neither SRC nor EW29Ch had detectable affinity for fully α2,3- (**602**, **705**, **706**, **710** and **712**) and α2,8-sialylated glycans (**707** and **713**). The essential aspects of the investigated features are expressed in a radar chart (Fig. 7C and D).

Though showing affinity for both α2-6Sia and Gal, it is still possible that SRC could have a single sugar-binding site. To examine this point, we carried out a competition assay using FAC. When α2,6-sialylated bi-antennary *N*-glycan (**503**) was eluted as a fluorescent probe in the presence of a competitor saccharide, lactose,

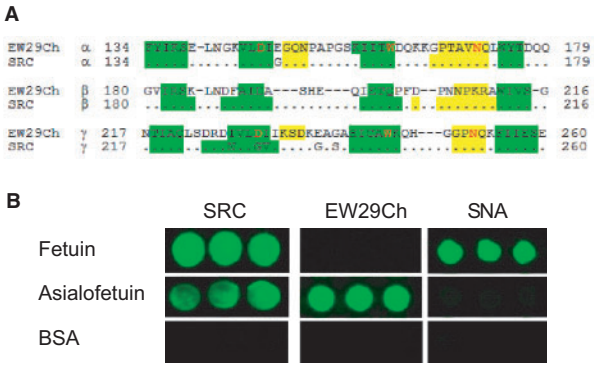

Fig. 6. **Overall characteristics of SRC.** (**A**) Sequence alignment of SRC and EW29Ch. Residues corresponding to 3_{10}-helices and β-strands are indicated in yellow and green, respectively. Sugar-binding residues of EW29Ch identified thus far are indicated with a red letter. Identical residues in SRC are indicated with a dot. (**B**) Evaluation by evanescent-field fluorescent-assisted lectin microarray using Cy3-labelled fetuin, asialofetuin and BSA (negative control). For comparison, SNA (Sia-binding) were also assayed as a positive control.

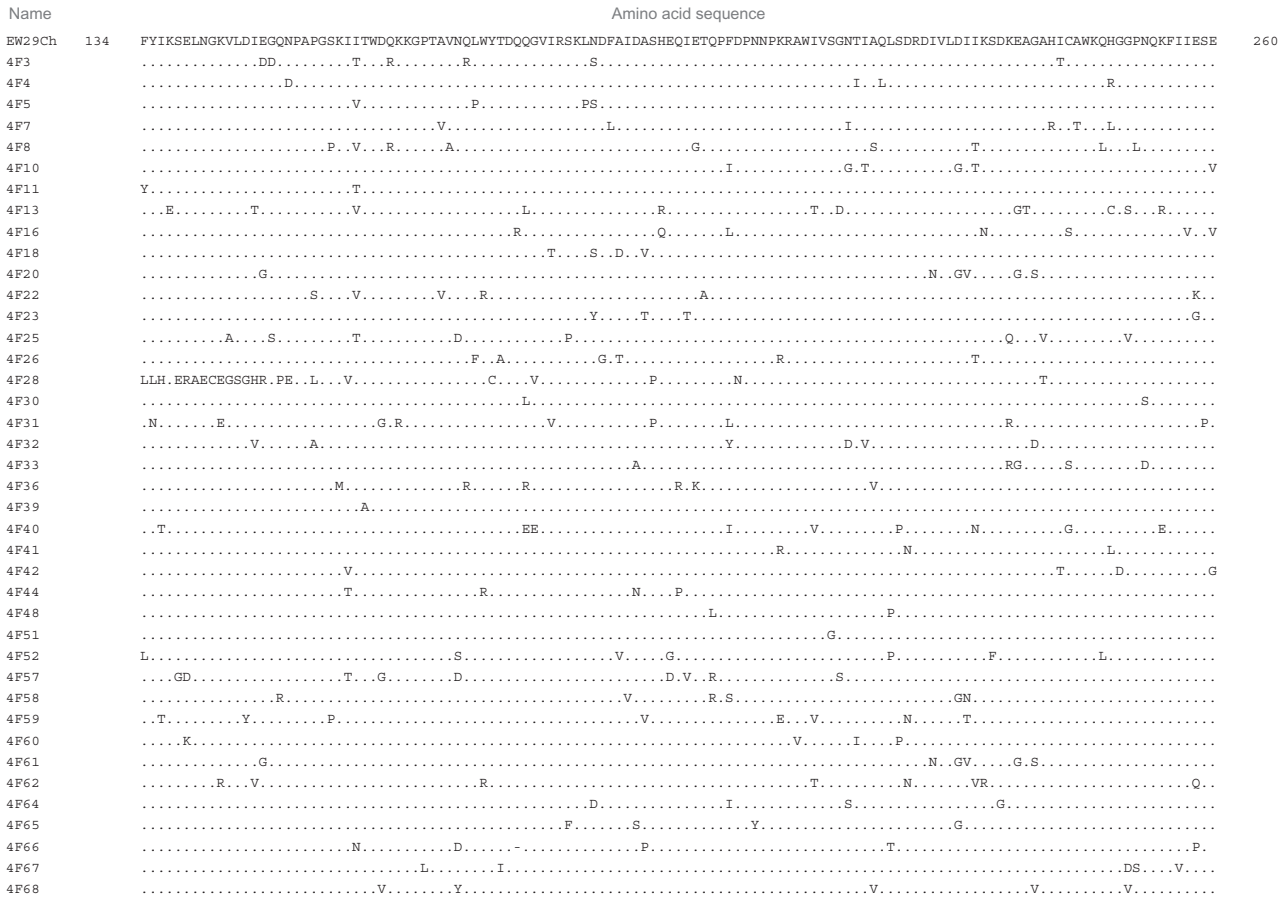

Fig. 5. **Sequence alignment of parental EW29Ch and molecular-evolved 40 mutants after the fourth selection cycle with fetuin–agarose.** Identical residues are indicated with a dot, whereas deletional ones are indicated by a hyphen.

the retardation was completely abolished (Fig. 8A). When fetuin was eluted, the retardation (monitored by fluorescence due to Trp) was also canceled by the presence of lactose (Fig. 8B). This result clearly demonstrated that SRC had a single binding pocket responsible for both lactose and α2,6-sialylated glycans. This conclusion is consistent with the observation that SRC had no hemagglutinating activity (data not shown). When compared with carbohydrate-binding properties of naturally occurring Sia-binding lectins, SNA and SSA, belonging to the R-type lectin family, the obtained SRC showed rather moderate binding affinity for α2,6-sialylated glycans and lowered but significant affinity for their asialoforms (unpublished data). In this context, SRC is considered to be an intermediate between original Gal-binding and evolved Sia-binding R-type lectins.

Structural Determination of SRC—To elucidate a molecular-evolutionary mechanism for SRC, we

Table 2. **Specifications of SRC- and EW29Ch-immobilized columns used in this study.**

Name	Immobilized (mg/ml)	B_t (nmol)	K_d (μM)
SRC	6.7	1.1	57 (pNP-β-6′SiaLac[a])
EW29Ch	1.9	0.16	18 (pNP-β-lactose)

[a]6′SiaLac, 6′-sialyllactose.

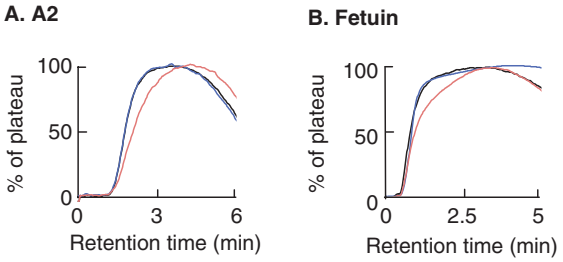

Fig. 8. **FAC competition assay.** Elution profiles obtained for the SRC-column are shown, where pyridylaminated (PA) α2,6-sialylated bi-antennary *N*-glycan (**503** in Fig. 3) (**A**) and glycoprotein fetuin (B) are eluted either in the absence (red) or presence (blue) of free lactose [10 mM in (A) and 200 mM in (**B**), respectively]. PA-oligosaccharide **503** and fetuin were detected by fluorescence at 310/380 and 285/350 nm, respectively.

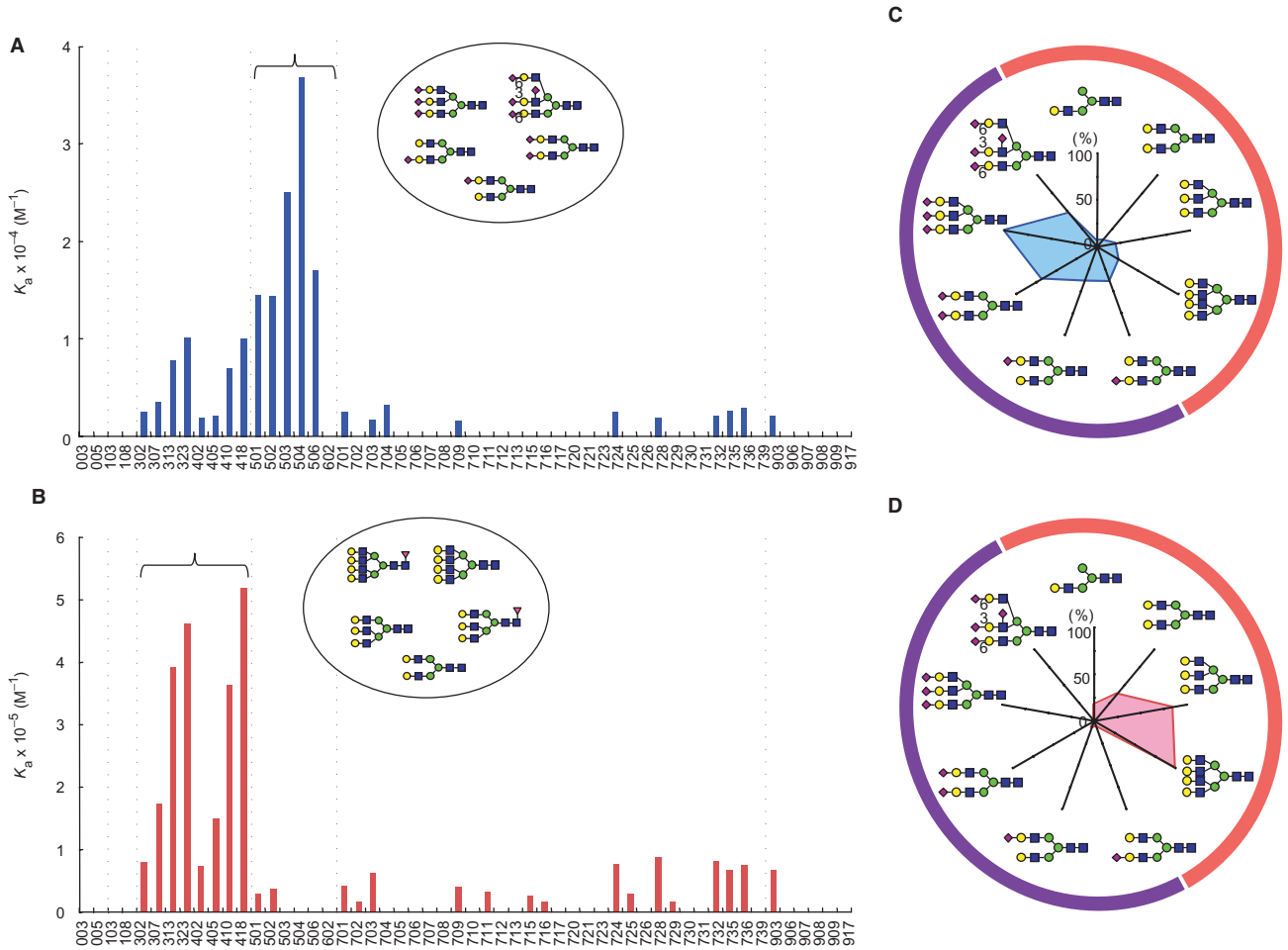

Fig. 7. **Comparison of the detailed sugar-binding specificities of SRC and EW29Ch by FAC.** Bar graph representation of K_a of SRC (A) and EW29Ch (B) for a total of 54 glycans is made for systematic comparison. Arabic numbers at the bottom of graphs correspond to sugar numbers indicated in Fig. 3. Radar chart representation of the binding pattern of SRC (C) and EW29Ch (D). Only selected members of either galactosylated- or sialylated *N*-glycans are depicted. Each radiation axis of the charts represents relative affinity (%), where the glycan showing the maximal affinity is taken as 100%.

performed X-ray crystallographic analysis for both the free and complex forms of SRC. For the latter, lactose and 6′-sialyllactose were used as a ligand (for crystal parameters and refinement statistics, see Table 1). The overall structure of SRC was identical to that of EW29Ch, *i.e*. β-trefoil fold (Fig. 9A and B) (Suzuki, R. *et al*., unpublished observations). However, a significant conformational change was observed in a large loop region of the subdomain γ (Fig. 9C). Since this area contains three mutations (D230G, I231V and E237G), these may have induced a dramatic flip of the side chains of W245 and I232, which in turn caused a dynamic shift (*i.e*. sharper hairpin formation) of the loop region containing E237G. This conformational change, together with an additional mutation, G239S, evolved a novel function, the recognition of α2-6Sia (described later). On the other hand, an electron density map corresponding to lactose was observed only in subdomain α of SRC (Fig. 9A). In this subdomain, the binding mechanism was essentially the same as that of EW29Ch (for detailed interactions, see Table 3).

The structure of SRC in complex with 6′-sialyllactose is shown in Fig. 9D. As in the case of SRC/lactose complex, an electron density map corresponding to the bound 6′-sialyllactose was observed only in the subdomain α of SRC (for detailed interactions, see Table 4). Notably, the manner of binding to the lactose unit of 6′-sialyllactose was identical to the binding to lactose (Fig. 9E). Importantly, additional interactions with α2-6Sia are depicted in SRC: they consist of two hydrogen bonds via G148 in subdomain α and S239 in the extended loop of subdomain γ (described earlier). In conclusion, SRC bears a single-binding site formed by subdomain α and the extended loop of subdomain γ, which has dual binding ability to α2-6Sia, and to a lesser extent, Gal.

DISCUSSION

A recent trend in molecular evolutionary engineering has been to focus mutation on local regions, such as complementarity-determining regions (CDR) in antibodies, substrate-binding sites in enzymes and loop regions of carbohydrate-recognition proteins (*41–43*). This type of approach, *i.e*. targeted mutagenesis, is represented by saturation mutagenesis and point mutagenesis based on modelling studies. In this report, we describe an alternative approach, *i.e*. random mutagenesis on an entire region, under the concept natural evolution-mimicry. Starting from a Gal-binding scaffold protein (EW29Ch), a novel Sia-binding lectin was successfully obtained by error-prone PCR and the reinforced ribosome display system. The resulting SRC preserved the practical advantages of high productivity in *E. coli*, and ease in purification using a conventional lactose-column. Both features apparently derived from the parental EW29Ch. Both lectin microarray and FAC analyses revealed specific SRC affinity for Sia. X-ray crystallographic analysis of SRC/sugar complexes demonstrated that SRC recognized 6′-sialyllactose with the conserved lactose-binding pocket (subdomain α) together with additional contacts in a distant region (γ). This cannot be easily attained by targeted mutagenesis.

As far as we know, there has been no report about three-dimensional structure of a Sia-binding lectin belonging to the R-type lectin family. Therefore, our present data will also be helpful to understand molecular recognition mechanisms of natural Sia-binding lectins, *e.g*. SNA, SSA and PSL, belonging to the same family (*17–21*).

Novel lectins, including SRC, produced by the above strategies should contribute, in various ways, to structural profiling of sialoglycoproteins, *i.e*. sialoglycomics. These include (i) lectin microarray analysis, (ii) FAC profiling, (iii) flow cytometry analysis and (iv) enrichment of sialoglycoproteins (*6–9*, *31*). The present method will also be applied to create other Sia-binding lectins, *e.g*. those specific for α2-3/α2-8Sia glycosides, which form critical biomarkers.

Supplementary data are available at JB online.

The authors thank S. Nakamura-Tsuruta, N. Uchiyama, J. Kominami and Y. Itakura, AIST, for critical discussion and technical supports about FAC and lectin microarray analyses. We also thank J. Iwaki, AIST, for a helpful comment about the X-ray crystallographic study. This work was supported in part by New Energy and Industrial Technology Development Organization (NEDO) in Japan.

Table 3. **Interaction between lactose and SRC.**

Ligand atom	Protein atom	Distance (Å)	Interaction
Gal O2	Q150 Nε2	3.0	Hydrogen bond
Gal O2	K164 Nζ	2.9	Hydrogen bond
Gal O3	D146 Oδ2	2.4	Hydrogen bond
Gal O3	N171 Nδ2	3.1	Hydrogen bond
Gal O4	D146 Oδ1	2.7	Hydrogen bond
Gal O4	G149 N	2.8	Hydrogen bond
Gal C3	W161 Cδ1	3.9	Stacking interaction
Gal C4	W161 Cε2	3.7	Stacking interaction
Gal C4	W161 Cδ2	3.6	Stacking interaction
Gal C4	W161 Cγ	3.9	Stacking interaction
Gal C5	W161 Cζ2	3.7	Stacking interaction
Gal C5	W161 Cη2	3.7	Stacking interaction
Gal C6	W161 Cε3	3.7	Stacking interaction
Gal C6	W161 Cζ3	3.5	Stacking interaction

Table 4. **Interaction between 6′-sialyllactose and SRC.**

Ligand atom	Protein atom	Distance (Å)	Interaction
Gal O2	Q150 Nε2	2.6	Hydrogen bond
Gal O2	K164 Nζ	2.9	Hydrogen bond
Gal O3	K164 Nζ	2.9	Hydrogen bond
Gal O3	D146 Oδ2	2.4	Hydrogen bond
Gal O3	N171 Nδ2	2.9	Hydrogen bond
Gal O4	D146 Oδ1	2.7	Hydrogen bond
Gal O4	G149 N	2.9	Hydrogen bond
Sia O1	S239 Oγ	2.7	Hydrogen bond
Sia O9	G148 O	2.8	Hydrogen bond
Gal C3	W161 Cδ1	3.9	Stacking interaction
Gal C4	W161 Cδ2	3.7	Stacking interaction
Gal C4	W161 Cγ	3.7	Stacking interaction
Gal C5	W161 Cε2	3.9	Stacking interaction
Gal C6	W161 Cε3	3.6	Stacking interaction
Gal C6	W161 Cζ3	3.7	Stacking interaction

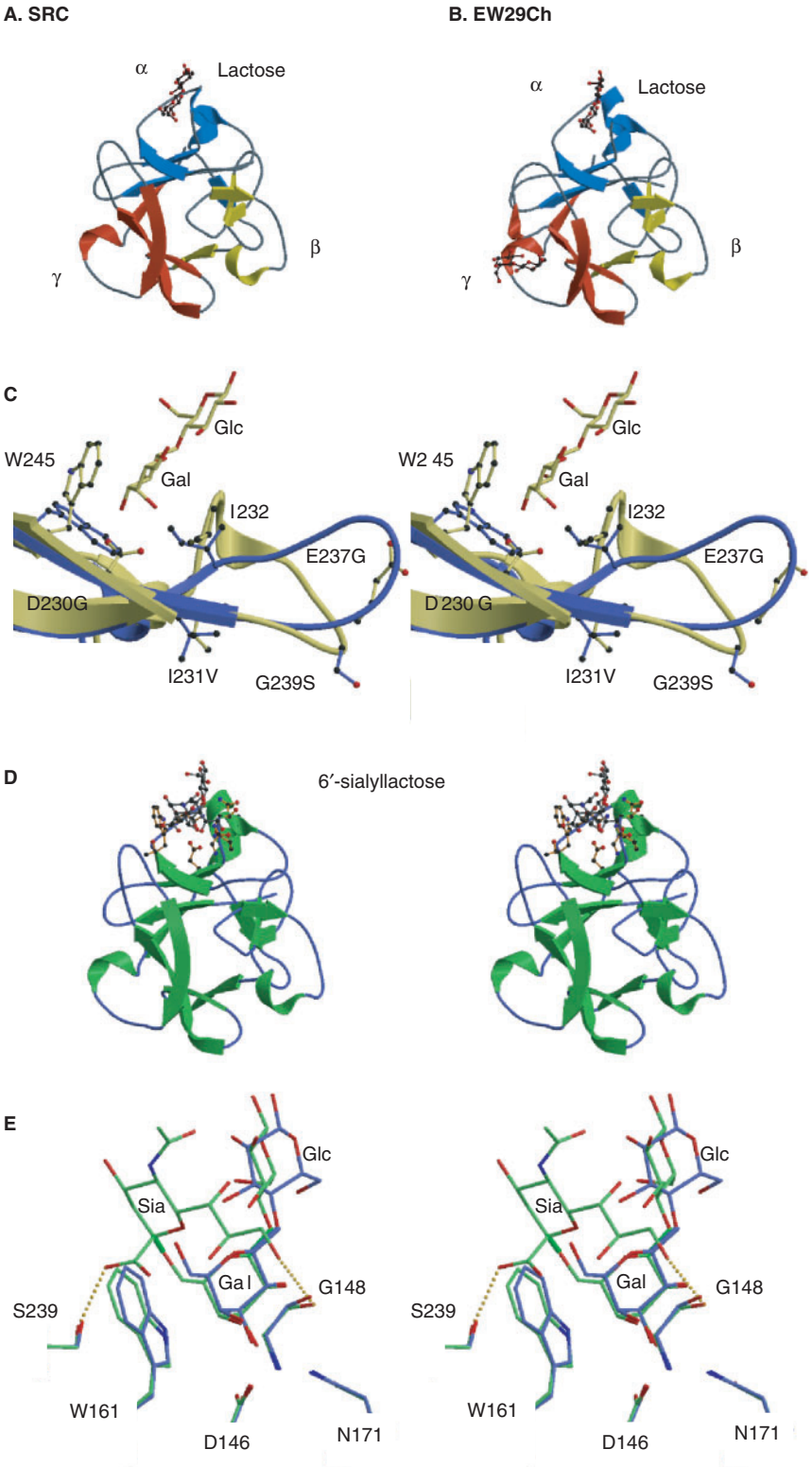

Fig. 9. **X-ray crystal structures of SRC and EW29Ch.** Overall structures of SRC (A) and EW29Ch (B) in complex with lactose. Subdomains α, β and γ are indicated in blue, yellow and red, respectively. (C) Stereoview of the binding feature of subdomain γ of SRC (blue), which is superimposed with EW29Ch (yellow). (D) Stereoview of the overall structure of SRC in complex with 6′-sialyllactose. (E) Stereoview of SRC in complex with 6′-sialyllactose (green), which is superimposed by that of lactose (blue). Only subdomain α is shown. The figures were drawn with the MOLSCRIPT and Raster3D programs (*44*, *45*).

REFERENCES

1. Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G., and Marth, J. (1999) *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
2. Sillanaukee, P., Pönniö, M., and Jääskeläinen, I.P. (1999) Occurrence of sialic acids in healthy humans and different disorders. *Eur. J. Clin. Invest.* **29**, 413–425
3. Kannagi, R. (2002) Regulatory roles of carbohydrate ligands for selectins in the homing of lymphocytes. *Curr. Opin. Struct. Biol.* **12**, 599–608
4. Kannagi, R., Cochran, N.A., Ishigami, F., Hakomori, S., Andrews, P.W., Knowles, B.B., and Solter, D. (1983) Stage-specific embryonic antigens (SSEA-3 and -4) are epitopes of a unique globo-series ganglioside isolated from human teratocarcinoma cells. *EMBO J.* **2**, 2355–2361
5. Kameyama, A., Kikuchi, N., Nakaya, S., Ito, H., Sato, T., Shikanai, T., Takahashi, Y., Takahashi, K., and Narimatsu, H. (2005) A strategy for identification of oligosaccharide structures using observational multistage mass spectral library. *Anal. Chem.* **77**, 4719–4725
6. Hirabayashi, J. (2004) Lectin-based structural glycomics: glycoproteomics and glycan profiling. *Glycoconj. J.* **21**, 35–40
7. Hirabayashi, J., Hayama, K., Kaji, H., Isobe, T., and Kasai, K. (2002) Affinity capturing and gene assignment of soluble glycoproteins produced by the nematode *Caenorhabditis elegans*. *J. Biochem. (Tokyo)* **132**, 103–114
8. 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
9. 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
10. Pilobello, K.T., Krishnamoorthy, L., Slawek, D., and Mahal, L.K. (2005) Development of a lectin microarray for the rapid analysis of protein glycopatterns. *Chem. Bio. Chem.* **6**, 985–989
11. Zheng, T., Peelen, D., and Smith, L.M. (2005) Lectin arrays for profiling cell surface carbohydrate expression. *J. Am. Chem. Soc.* **127**, 9982–9983
12. Ebe, Y., Kuno, A., Uchiyama, N., Koseki-Kuno, S., Yamada, M., Sato, T., Narimatsu, H., and Hirabayashi, J. (2006) Application of lectin microarray to crude samples: differential glycan profiling of Lec mutants. *J. Biochem. (Tokyo)* **139**, 323–327
13. Koshi, Y., Nakata, E., Yamane, H., and Hamachi, I. (2006) A fluorescent lectin array using supramolecular hydrogel for simple detection and pattern profiling for various glycoconjugates. *J. Am. Chem. Soc.* **128**, 10413–10422
14. Hsu, K.L., Pilobello, K.T., and Mahal, L.K. (2006) Analyzing the dynamic bacterial glycome with a lectin microarray approach. *Nat. Chem. Biol.* **2**, 153–157
15. Lee, M.R., Park, S., and Shin, I. (2006) Protein microarrays to study carbohydrate-recognition events. *Bioorg. Med. Chem. Lett.* **16**, 5132–5135
16. Hirabayashi, J. (2004) On the origin of glycome and saccharide recognition. *Trends Glycosci. Glycotechnol.* **16**, 63–85
17. 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(α2-6)Gal/GalNAc sequence. *J. Biol. Chem.* **262**, 1596–1601
18. Van Damme, E.J., Barre, A., Rouge, P., Van Leuven, F., and Peumans, W.J. (1996) The NeuAc(α-2,6)-Gal/GalNAc-binding lectin from elderberry (*Sambucus nigra*) bark, a type-2 ribosome-inactivating protein with an unusual specificity and structure. *Eur. J. Biochem.* **235**, 128–137
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. Mo, H., Winter, H.C., and Goldstein, I.J. (2000) Purification and characterization of a Neu5Acα2-6Galβ1-4Glc/GlcNAc-specific lectin from the fruiting body of the polypore mushroom *Polyporus squamosus*. *J. Biol. Chem.* **275**, 10623–10629
21. Tateno, H., Winter, H.C., and Goldstein, I.J. (2004) Cloning, expression in *Escherichia coli* and characterization of the recombinant Neu5Acα2,6Galβ1,4GlcNAc-specific high-affinity lectin and its mutants from the mushroom *Polyporus squamosus*. *Biochem. J.* **382**, 667–675
22. Imberty, A., Gautier, C., Lescar, J., Perez, S., Wyns, L., and Loris, R. (2000) An unusual carbohydrate binding site revealed by the structures of two *Maackia amurensis* lectins complexed with sialic acid-containing oligosaccharides. *J. Biol. Chem.* **275**, 17541–17548
23. Hirabayashi, J., Hashidate, T., Arata, Y., Nishi, N., Nakamura, T., Hirashima, M., Urashima, T., Oka, T., Futai, M., Muller, W.E., Yagi, F., and Kasai, K. (2002) Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. *Biochim. Biophys. Acta* **1572**, 232–254
24. Leppanen, A., Stowell, S., Blixt, O., and Cummings, R.D. (2005) Dimeric galectin-1 binds with high affinity to α2,3-sialylated and non-sialylated terminal *N*-acetyllactosamine units on surface-bound extended glycans. *J. Biol. Chem.* **280**, 5549–5562
25. Hanes, J. and Plückthun, A. (1997) *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc. Natl. Acad. Sci. USA* **94**, 4937–4942
26. Hirabayashi, J., Dutta, S.K., and Kasai, K. (1998) Novel galactose-binding proteins in Annelida. *J. Biol. Chem.* **273**, 14450–14460
27. Ito, S., Kuno, A., Suzuki, R., Kaneko, S., Kawabata, Y., Kusakabe, I., and Hasegawa, T. (2004) Rational affinity purification of native *Streptomyces* family 10 xylanase. *J. Biotechnol.* **110**, 137–142
28. Cadwell, R.C. and Joyce, G.F. (1992) Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* **2**, 28–33
29. Hanes, J., Jermutus, L., and Plückthun, A. (2000) Selecting and evolving functional proteins *in vitro* by ribosome display. *Methods Enzymol.* **328**, 404–430
30. Suzuki, R., Fujimoto, Z., Kuno, A., Hirabayashi, J., Kasai, K., and Hasegawa, T. (2004) Crystallization and preliminary X-ray crystallographic studies of the C-terminal domain of galactose-binding lectin EW29 from the earthworm *Lumbricus terrestris*. *Acta Crystallogr. D* **60**, 1895–1896
31. 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
32. Arata, Y., Hirabayashi, J., and Kasai, K. (2001) Application of reinforced frontal affinity chromatography and advanced processing procedure to the study of the binding property of a *Caenorhabditis elegans* galectin. *J. Chromatogr. A* **905**, 337–343
33. Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326
34. Matthews, B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497
35. Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022–1025

36. Collaborative Computational Project, Number 4.(1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763
37. Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. (1998) *Crystallography & NMR system*: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921
38. Engh, R.A. and Huber, R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A* **47**, 392–400
39. McRee, D.E. (1999) XtalView/Xfit-A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165
40. Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486
41. Hoogenboom, H.R. (2005) Selecting and screening recombinant antibody libraries. *Nat. Biotechnol.* **23**, 1105–1116
42. Bloom, J.D., Meyer, M.M., Meinhold, P., Otey, C.R., MacMillan, D., and Arnold, F.H. (2005) Evolving strategies for enzyme engineering. *Curr. Opin. Struct. Biol.* **15**, 447–452
43. Yim, M., Ono, T., and Irimura, T. (2001) Mutated plant lectin library useful to identify different cells. *Proc. Natl. Acad. Sci. USA* **98**, 2222–2225
44. Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950
45. Merritt, E.A. and Bacon, D.J. (1997) Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524