J. Biochem. **142,** 393–401 (2007)
doi:10.1093/jb/mvm146

Elderberry Bark Lectins Evolved to Recognize Neu5Acα2,6Gal/GalNAc Sequence from a Gal/GalNAc Binding Lectin Through the Substitution of Amino-Acid Residues Critical for the Binding to Sialic Acid

Hanae Kaku[1,3,*], Hiroki Kaneko[2,*], Naoto Minamihara[3,4,*], Kazumichi Iwata[3], Elizabeth T. Jordan[3], Maria A. Rojo[3], Naoko Minami-Ishii[3], Eiichi Minami[3], Shigeru Hisajima[4] and Naoto Shibuya[1,3,†]

[1]Department of Life Sciences, Meiji University, Kawasaki, Kanagawa 214-8571; [2]Department of Integrated Sciences in Physics and Biology, Nihon University, Tokyo 156-8550; [3]Department of Biochemistry, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602; and [4]Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan

Received April 20, 2007; accepted June 29, 2007; published online July 23, 2007

Bark lectins from the elderberry plants belonging to the genus *Sambucus* specifically bind to Neu5Acα2,6Gal/GalNAc sequence and have long been used for the analysis of sialoglycoconjugates that play important roles in many biological phenomena. However, molecular basis of such a unique carbohydrate binding specificity has not been understood. To answer these questions, we tried to identify the amino-acid residues in the Japanese elderberry bark lectin, *Sambucus sieboldiana* agglutinin that enabled the lectin to recognize sialic acid by using *in silico* docking simulation and site-directed mutagenesis. These studies showed that three amino-acid residues, S_{197}, A_{233} and Q_{234}, in the C-terminal subdomain of SSA-B chain are critical for the binding to the sialic acid in Neu5Acα2,6Gal/GalNAc sequence. Replacement of one of these residues to the one in the corresponding position of ricin B-chain completely abolished the binding to a sialoglycoprotein, fetuin. Conserved presence of these amino acid residues in the corresponding sequences of two other elderberry lectins with similar binding specificity further supported the conclusion. These findings indicated that the replacement of the corresponding amino-acid residues in a putative Gal/GalNAc-specific ancestral lectin to these amino-acid residues generated the unique Neu5Acα2,6Gal/GalNAc-specific elderberry lectins in the course of molecular evolution.

Key words: docking simulation, elderberry, lectin, *Sambucus sieboldiana*, sialic acid.

Abbreviations: PBS, phosphate buffered saline; PDB, protein data bank; RIP, ribosome-inactivating protein; SNA, *Sambucus nigra* agglutinin; SSA, *Sambucus sieboldiana* agglutinin.

Plant lectins with defined carbohydrate binding specificities have been used as invaluable tools for the detection, fractionation and isolation of glycoconjugates (*1*). Compared to the popular use of plant lectins in glycobiology and related fields, their biological roles are not well understood, though their involvement in the plant defense systems has been indicated. As the carbohydrate binding specificity of plant lectin is the base of both its application as well as the presumed biological functions, it is important to clarify the molecular basis of the specific recognition of carbohydrates by these lectins.

Several bark lectins from the elderberry plants belonging to the genus *Sambucus* specifically bind to Neu5Acα2,6Gal/GalNAc sequence (*2*, *3*) and have long been used for the analysis of sialoglycoconjugates (*4*–*7*) that play important roles in many biological as well as pathological phenomena (*8*). A fruit lectin from the elderberry was also shown to bind to sialoglycoprotein (*9*). The binding specificity of these elderberry lectins is quite unique and so far no other lectin with similar binding specificity has been isolated from plants, except *Maackia amurensis* lectin that recognizes α2,3-linked sialylated oligosaccharides (*10*, *11*).

The elderberry bark lectins are tetrameric glycoproteins consisting of two types of subunits, one with a carbohydrate binding site and another one without it. The cDNA cloning as well as molecular modeling studies revealed that each subunit of elderberry bark lectins has highly homologous structure to the galactose-specific lectin subunit (B chain) and ribosome- inactivating subunit (A chain) of type 2 ribosome-inactivating proteins (RIPs), such as ricin and abrin (*12*, *13*), indicating a close evolutionary relationship between these proteins. While the lectin subunit of ricin B chain preferably binds to Gal/GalNAc, the corresponding subunit of *Sambucus sieboldiana* agglutinin (SSA) specifically binds to Neu5Acα2,6Gal/GalNAc sequence, as described.

*The first three authors contributed equally to this work.
†To whom correspondence should be addressed. Tel: +81-44-934-7039, E-mail: shibuya@isc.meiji.ac.jp

 © 2007 The Japanese Biochemical Society.

Downloaded from http://jb.oxfordjournals.org/ at University of Science and Technology of China on September 28, 2012

Scatchard analysis in the binding studies (*2*) as well as the establishment of monomeric-monovalent derivative of SSA (*5*) showed that the carbohydrate-binding subunit of elderberry lectins carries one carbohydrate binding site, whereas the ricin B chain carries two carbohydrate binding sites with different binding characteristics (*14*, *15*).

It has not been understood how the elderberry bark lectins obtained such a unique carbohydrate binding specificity and what could be the outcome of such molecular evolution for these plants themselves. Although Chen *et al*. (*16*) reported that the D_{231} and N_{48} of SNA are involved in the binding to galactose and fetuin from the comparison of the amino-acid sequence with a structurally related elderberry protein without lectin activity as well as from the site-directed mutagenesis, these results did not explain how the elderberry bark lectins obtained the unique binding specificity to Neu5Acα2,6Gal/GalNAc sequence. To answer these questions, we tried to find out the critical changes in the structure of Japanese elderberry bark lectin, SSA that enabled the lectin to recognize sialic acid residue in Neu5Acα2,6Gal/GalNAc sequence. In this article, we describe the heterologous expression of the carbohydrate binding domain of SSA-B subunit in *Escherichia coli* and identification of amino-acid residues critical for the binding to Neu5Acα2,6Gal/GalNAc sequence, especially to the sialic acid residue, by using *in silico* docking simulation and site-directed mutagenesis of the carbohydrate binding domain. Possible course of molecular evolution from Gal/GalNAc binding lectin to Neu5Ac α2,6Gal/GalNAc lectin and its outcome for the plant are also discussed.

MATERIALS AND METHODS

Preparation of Japanese Elderberry Bark Lectin and its Antibody—The Neu5Acα2,6Gal/GalNAc-specific bark lectin (SSA) was purified from the extract of the twigs of Japanese elderberry (*Sambucus sieboldiana*) by affinity chromatography on fetuin–agarose as previously reported (*17*). Rabbit anti-SSA antibody was also raised and purified as described (*17*).

Construction of Expression Vectors—The DNA fragment corresponding to SSA-B chain was amplified by PCR using a forward primer with *Nde* I site at 5′ end [5′-CGCAT ATGGGGGGCGAGTACGAAAA AGTG-3′] and a reverse primer with *Bam*H I site at 3′ end [5′-CGGGATCCC TAAGCTGGCTGTGTG GTA-3′, designated as primer A], and the plasmid for the whole SSA cDNA as a template (*12*). The PCR product was isolated from low melting agarose (0.7% sea plaque agarose) gel, ligated into pCR II vector (Invitrogen Co., Carlsbad, USA) using TaKaRa Ligation Kit ver.2 (Takara-Bio Inc., Sigma, Japan) and transformed into *E. coli* JM109. The sequence of inserted DNA of a positive colony was confirmed by an ALF DNA sequencer (Amersham Biosciences, Piscataway, USA).

The plasmid encoding SSA-B chain was digested with restriction endonucleases *Nde* I and *Bam*H I. The obtained fragment was subcloned into an expression vector pET-16b that adds a poly-histidine tag (His-tag) to the N-terminal of the expressed protein (Novagen, Darmstadt, Germany) and transformed into *E. coli* BL21 (DE3) according to the manufacturer's instruction.

The DNA fragment for SSA-B chain amplified with the pCR II plasmid was further divided into two parts for N-terminal region (SSA-BN, 1–348 bp) and C-terminal region (SSA-BC, 349–786 bp) by the treatment with *Nde* I/*Xho* I and *Xho* I/*Bam*H I, respectively. Both fragments were subcloned again into the pET-16b expression vector. The plasmid DNA for SSA-BC domain was also subcloned into pET-5a (Novagen, Darmstadt, Germany) vector and transformed into *E. coli* BL21 (DE3) to express SSA-BC domain without the His-tag.

Site-directed Mutagenesis of SSA-BC Domain—PCR-based site-directed mutagenesis was performed as follows. Two sets of PCR were used to replace serine197 of SSA-BC domain to glutamic acid ($S_{197}E$). The primer sets used were, (i) the primer for *Nde*I site at 5′ end [5′-CGCATATGCTCGAGAAAAATA TCCATGC CG-3′, designated as primer B] and $S_{197}E1$ antisense primer [5′-CCACGAGCCCGAAGATC TTATCG TCA] and (ii) $S_{197}E2$ sense primer [5′-TGACG ATAAGATCTTC GGGCTCGTGG] and primer A, respectively. PCR was carried out by using *EX Taq* polymerase (Takara-Bio Inc., Sigma, Japan) and the plasmid for the SSA-BC domain as a template. Both PCR products were purified by agarose gel electrophoresis, added together into the PCR buffer and heated for 3 min at 94°C. The solution was cooled down to 37°C over 1 h, kept for 15 min, added with *EX Taq* polymerase and primers A and B, then subjected to PCR amplification. The PCR product was ligated into pCRII vector and transformed into *E. coli* JM109. After confirming the sequence, the pcr amplicon was treated with *Nde* I and *Bam*H I, and the resulting fragments were subcloned to pET 5a and transformed into *E. coli* BL21(DE3).

The other four site-directed mutagenesis, $A_{233}R$, $Q_{234}A$, $K_{241}Q$ and $D_{198}T$ in the SSA-BC domain were also carried out similarly by using the following primer sets. Antisense primers; $A_{233}R1$ [5′-TATGGACGTTAGACA ACACGATTCT-3′], $Q_{234}A1$ [5′-TATGGACGTTGCAGCAC ACGATATCT-3′], $K_{241}Q1$ [CTCTCTTCGACAAATCATTC TCT], $D_{198}T1$ [CCACGAGCCCAGTACTCTTATCGTCA]. Sense primers; $A_{233}R2$ [5′-GAGTATCGTGTTGTCTAAC GTCCATA-3′], $Q_{234}A2$ [5′-AGATATCGTGTGCT GCAAC GTCCATA-3′], $K_{241}Q2$[AGAGAATGATTTGTCGAAGAG AG], $D_{198}T2$ [TGACGATAAGAG TACTGGGCTCGT GG].

Purification and Refolding of Recombinant Proteins—The *E. coli* for the expression of recombinant protein was pre-cultured with 5 ml LB-Ampicilin (LB-Amp) medium overnight with 250 rpm at 37°C and then transferred to 250 ml of the same buffer. IPTG was added to the medium to make a final 0.5 mM when the turbidity at 600 nm reached 0.6–0.8 and further incubated for 3 h. The *E. coli* cells were harvested by centrifugation (10,000 rpm) for 5 min at 4°C. The harvested cells were washed with PBS containing 1 mM PMSF and homogenized with a French Press. The inclusion body fraction was recovered from the homogenate by ultracentrifugation (36,000 rpm) for 1 h at 4°C.

The His-tagged recombinant proteins (His-rSSA-B, His-rSSA-BN and His-rSSA-BC) were purified by

Downloaded from http://jb.oxfordjournals.org/ at University of Science and Technology of China on September 28, 2012

affinity chromatography on Ni-Sepharose 6B (Novagen, Darmstadt, Germany). Briefly, the inclusion body fraction from a 250 ml culture was solubilized with 40 ml of the binding buffer (5 mM imidazole, 0.5 M NaCl in 20 mM Tris-HCl, pH 7.9) containing 6 M urea and ultracentrifuged. The solubilized protein was applied to a Ni-Sepharose 6B column (2.5 ml) pre-equilibrated with the same buffer. After washed with the binding buffer, the bound protein was eluted with the elution buffer (6 M urea, 1 M imidazole, 0.5 M NaCl in the 20 mM Tris–HCl, pH 7.9). The eluted protein fraction was applied to a gel filtration column, PD-25G, Amersham Biosciences, Piscataway, USA, equilibrated with the denaturing buffer [8 M urea, 10 mM glutathione reduced form (GSH) in 20 mM Tris–HCl, pH 8.8] to exchange buffer.

Refolding of denatured His-tagged proteins was carried out with the combination of reduced (GSH) and oxidized form (GSSG) of glutathione. The purified His-tagged protein in the denaturing buffer was kept for 5 min at 65°C, then immediately diluted 40 times with the refolding buffer (8 M urea, 10 mM GSH, 4 mM GSSG in 20 mM Tris–HCl, pH 8.8). The mixture was kept for 4 h at room temperature or overnight at 4°C, then dialysed against PBS overnight.

Recombinant proteins without His-tag, such as rSSA-BC domain and those with the site-directed mutagenesis were also subjected to denaturation/renaturation as described earlier. The renatured rSSA-BC protein was purified by using affinity chromatography on fetuin–agarose (*17*). The rSSA-BC protein that bound to the column was eluted with 0.5 M lactose in PBS, then subjected successively to gel filtration (PD-25G) and dialysis against PBS to remove lactose.

The mutagenized SSA-BC protein was applied to an anti-SSA antibody-Sepharose 4B column pre-equilibrated with PBS and the bound fraction was eluted with 0.17 M glycine–HCl (pH 2.3). The fractions were collected in the tubes containing one per tenth volume of 1 M Tris–hydroxymethylaminomethane for immediate neutralization.

Homology Modeling and Ligand Docking—Homology modelling of SSA was performed by using the following steps. (i) Homology search from Protein Data Bank (PDB); (ii) Sequence alignment with an algorithm (*18*) that is a variant of the Needleman and Wunsch algorithm (*19*) but improved for the accuracy of the alignment owing to the introduction of the concept of hydrophobic core scores (*20*); (iii) Loop searching and loop replacement when insertion/deletion sites are required; (iv) Side chain replacement; (v) Removal of the steric hindrances between side chains, in which 10,000 random conformational searches for each side chain were carried out by using random numbers which reproduce normal distribution of dihedral angles χ experimentally found in crystals (*21*); (vi) Energy minimization. A program package for protein engineering and drug design, BIOCES[E] (NEC Corp., Tokyo, Japan), was used for all the above molecular modelling. This package runs on an OCATANE2 (Silicon Graphics Inc., Mountain View, CA, USA).

Docking of ligands and molecular dynamics simulations of the complex models were carried out using Discover/Insight II (Accelrys Inc., San Diego, CA, USA) on an OCATANE2 and Amber on an Express5800/MD Server (NEC Corp., Tokyo, Japan), respectively.

Analysis of Carbohydrate Binding Properties by ELISA—Lectin activity was determined by enzyme-linked immunosorbent assay (ELISA) as reported previously (*5*). Briefly, a 96-well flat-bottom microtitre plate was coated with fetuin solution (1 μg/50 μl/well) in 0.1 M carbonate buffer (pH 8.5) and kept overnight at 4°C. The fetuin solution was then discarded, and the plate was incubated with 1% BSA in the same buffer (100 μl) for 1 h at 37°C to prevent non-specific binding. The wells were rinsed three times with PBS/0.05% Tween 20 (this rinsing was repeated after all subsequent steps). Then, 50 μl of PBS containing each recombinant protein (0.1% BSA/0.05% Tween 20) was added to each well and incubated for 1 h at 37°C. The protein bound to the coated fetuin was detected with the partially purified anti-SSA antibody (1/2,000) and HRP-labelled goat anti-rabbit IgG (1/2,000, Sigma Chemical Co., St Louis, MO, USA). After incubation with the substrate [2,2′-azino-bis(3- ethylbenzothiazoline-6-sulfonic acid) diammonium salt], the colour intensity at 415 nm was determined by a microplate reader (MTP-32, Corona Electric Co., Ltd., Japan).

Hapten inhibition experiments were carried out as follows. A solution of PBS–Tween–BSA containing varying amount of test sugars (50 μl) was added to each well of a fetuin-coated plate and incubated at 4°C for 30 min. Then, each lectin solution (100 ng/50 μl) was added and incubated for 1 h. Determination of the bound protein with the anti-SSA antibody was carried out as described previously.

Analysis of Carbohydrate Binding Properties with a Biosensor (Biacore)—Fetuin was dissolved in 10 mM sodium acetate buffer (pH 4.8) to make a 0.1 mg/ml solution and immobilized on a sensor tip for Biacore (Biacore AB Corp, Uppsala, Sweden) as described by the operation manual. Each one of the wild-type SSA (30 μg/ml), rSSA-BC (10 μg/ml), or mutagenized SSA-BC protein (10 μg/ml) was dissolved with 10 mM HEPES buffer (pH7.4) containing 0.15 M sodium chloride, 3.4 mM EDTA and 0.05% Tween 20 and reacted with the immobilized lectin using Biacore 1000. Kinetic parameters were calculated from the sensorgram using the program installed in the instrument.

RESULTS

Heterologous Expression and Characterization of Carbohydrate Binding Domain of SSA—For the detailed study of the carbohydrate recognition by SSA with molecular modelling as well as site-directed mutagenesis, it was necessary to establish a suitable heterologous expression system and also to know which subunit, A or B subunit, carries the carbohydrate binding site. Concerning to the carbohydrate binding subunit, it was confirmed that the B chain of SSA carries the binding site by the analysis of N-terminal amino-acid sequence of the subunit, prepared as a form of stable monomer (*5*), that was adsorbed to fetuin–agarose and recovered (data not shown).

Downloaded from http://jb.oxfordjournals.org/ at University of Science and Technology of China on September 28, 2012

For the heterologous expression, several eukaryotic systems such as yeast (*22*), *Xenopus oocytes* (*15*, *23*), COS cells (*24*, *25*), have been preferably used for the expression of ricin B chain. In this study, however, we tried to establish the conditions of the expression in *E. coli* as well as the recovery of soluble proteins carrying carbohydrate binding activity because of the advantage for the production of enough amount of recombinant protein. The B chain was first expressed as a poly-histidine-tagged protein (His-rSSA-B) for the ease of purification. As most of the expressed B chain was recovered in the form of inclusion body, we tried to establish a condition to recover the active, renatured protein by successive denaturation/renaturation treatment. We found that the conditions for the immediate dilution step after the denaturation by urea, *e.g.* the urea concentration after dilution and the ratio of oxidized/reduced form of glutathione in the renaturation solution, greatly affected the recovery of the active lectin, which was estimated by using the ELISA-type binding assay (*5*, *26*). The finding that the ratio of GSSG/GSH is very important for the refolding is understandable as the SSA-B chain contains 10 cysteine residues, making a numerous possibility of mismatched disulfide linkage formation. Under the optimal conditions we established, over 32% of the expressed protein was recovered as soluble protein with the binding activity (data not shown).

By using the similar approach, we then expressed N- and C-terminal domains of SSA-B chain separately (1–116 amino-acid residues for N-terminal domain, 117–262 amino-acid residues for C-terminal domain) and examined which domain carries the binding site (Fig. 1). As shown in Fig. 2, the expressed C-terminal domain (His-rSSA-BC) showed the binding to fetuin, whereas the N-terminal domain (His-rSSA-BN) did not bind, indicating the carbohydrate binding site resides in the C-terminal domain of the B subunit. As the His-rSSA-BN prepared in this experiment lacked some amino-acid residues corresponding to a part of γ-subdomain of ricin-like lectins, it seems impossible to exclude the presence of other carbohydrate binding site(s) in the N-terminal domain only from this experiment. However, combining the results with the previous finding that SSA carries two carbohydrate binding sites in the tetrameric molecule consisting of two A and two B subunits (*3*), the carbohydrate binding subunit, B-chain, should contain only one carbohydrate binding site in its C-terminal domain.

Analysis of the binding specificity of the expressed His-rSSA-BC protein resulted in an unexpected observation. On the contrary to SSA, which was specific to only 2,6-linked sialylated oligosaccharides/glycoconjugagtes, the interaction of the expressed His-rSSA-BC protein and fetuin was inhibited by both 2,3- and 2,6-linked sialyllactose (data not shown). As the expressed C-domain was tagged with poly-histidine sequence, we suspected that the addition of the poly-cationic group might have affected such change in the binding specificity. Thus, we expressed the C-terminal domain of the B subunit without the poly-histidine tag (rSSA-BC), followed by the affinity purification on a fetuin–agarose column (Fig. 3A). As shown in Fig. 3B and C, the rSSA-BC protein expressed without the

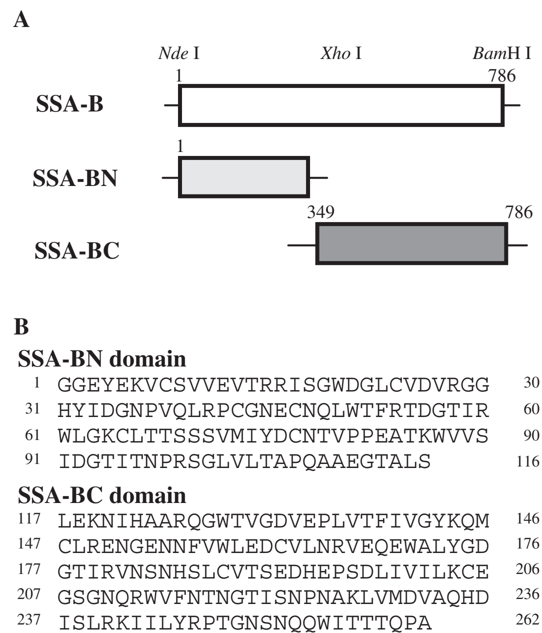

Fig. 1. **Structure of SSA-B subunit and its N-/C-terminal domains.** (A) Schematic representation of cDNA sequences encoding SSA-B subunit (SSA-B), its N-terminal domain (SSA-BN) and C-terminal domain (SSA-BC). (B) The amino-acid sequences of SSA-BN and SSA-BC.

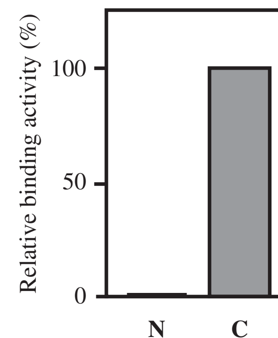

Fig. 2. **Binding activity of the N- and C-terminal domains of SSA-B subunits to fetuin.** Fetuin solution was coated on a 96-well flat bottom plate (1 μg/well). Binding activity of the N-terminal (N) and C-terminal (C) domains of SSA-B subunit (each 100 ng/well) to fetuin was analysed with the His-tagged proteins prepared as described in the experimental procedures.

poly-histidine tag showed the binding specificity towards Neu5Acα2,6 lactose as similar to the original SSA (Concentrations required for 50% inhibition were 0.048 and 0.032 mM for Neu5Acα2,6 lactose, and 11.2 and 9.3 mM for lactose, both for SSA and rSSA-BC, respectively). On the other hand, both of the binding of rSSA-BC and SSA to fetuin were not inhibited by Neu5Acα2,3 lactose, showing the poly-histidine tag really affected the binding specificity of His-rSSA-BC protein.

Analysis of the Amino-Acid Residues Critical for the Binding to Sialylated Oligosaccharides—To find proper candidates of the amino-acid residues involved in the binding to silaylated oligosaccharides, we performed

Downloaded from http://jb.oxfordjournals.org/ at University of Science and Technology of China on September 28, 2012

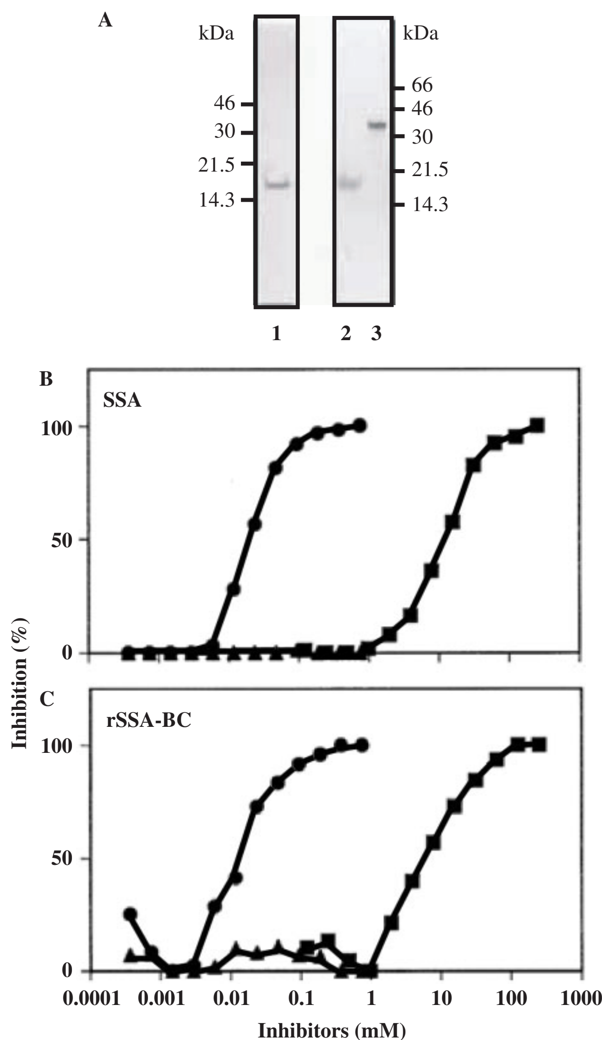

Fig. 3. **Binding specificity of the expressed rSSA-BC protein.** (A) The rSSA-BC protein without His-tag was expressed and purified as described in the experimental procedures. SDS-PAGE was carried out using a 15% acrylamide gel, followed by CBB staining (1, rSSA-BC, 6 μg) or Western blotting with anti-SSA antibody (2, rSSA-BC, 1 μg; 3, SSA, 0.2 μg). Binding specificity of SSA **(B)** and rSSA-BC protein **(C)** was analyzed by inhibition of the binding to fetuin with haptenic sugars. Varying amount of each sugar (50 μl) was added to each well of the fetuin-coated plate and incubated for 30 min at 4°C. SSA or rSSA-BC solution (100 ng/50 μl) was added to each well and incubated for 1 h. The bound protein was detected with anti-SSA antibody as described in experimental procedures. Filled circles, Neu5Acα2,6lactose; filled triangles, Neu5Acα2,3lactose; filled squares, lactose.

homology modelling as well as ligand docking studies. As the amino-acid sequence of the SSA-B chain showed a very high homology (46% identity over the entire sequence) with the B chain of ricin (PDB ID: 2AAI) (*12*, *14*), the B chain of ricin was selected as the most suitable template for homology modelling.

The model of the SSA–ligand complex was built as follows. We used the information on the crystal structure (PDB ID: 2CWG) (*27*) of wheat germ lectin complexed with an *O*-linked tetrasaccharide, Neu5Acα2,3Galβ1,3

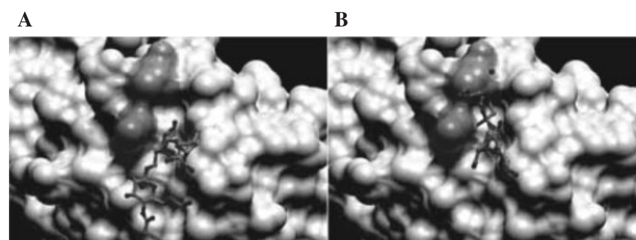

Fig. 4. **Neu5Acα2,6lactose (A) and Neu5Acα2,3lactose (B) in the binding cleft of SSA.** Connolly surface of SSA model is shown. Neu5Acα2,6lactose and Neu5Acα2,3lactose are displayed by coloured ball-and-stick in each. Y_{245} and T_{248}, which cause steric hindrance with the Neu5Acα2,3lactose, are shown in orange and pink, respectively.

(Neu5Ac α2,6) GalNAc-α-1-*O*-Thr, and also the B chain of ricin (PDB ID: 2AAI) (*14*) complexed with its ligand, lactose, to generate 3D structure of the SSA–ligand complex. The structure of the Neu5Acα2,6GalNAc, which is a part of the tetrasaccharide in 2CWG, was superimposed on the structure of the lactose in 2AAI. The superposition was performed by least square fitting of the common galactose frame, generating the complex structure of the B chain of ricin and Neu5Acα2,6lactose. Then the Neu5Acα2,6 lactose was docked into the binding site of SSA on a 3D display referring to the structure of the complex described earlier. The complex model was further refined by molecular dynamic calculation. We also tried to dock Neu5Acα2,3 lactose forcibly into the cleft of SSA by the similar approach. In this case, the structure of Neu5Acα2,3lactose was taken from the published crystal structure (PDB ID: 1WGC) of the complex between wheat germ lectin and Neu5Acα2,3lactose (*28*). As shown in Fig. 4, Neu5Acα2,6lactose was smoothly docked to the binding cleft of SSA but Neu5Acα2,3lactose did not fit to the cleft because of the steric hindrance between the sialic acid residue and the surrounding amino-acid residues such as T_{248} and Y_{245}. Glucosyl residue of Neu5Acα2,6lactose seemed to be extruded from the binding cleft, different from the galactosyl and sialic acid residues that are located within the binding cleft. This coincides with the previous observation with NMR analysis that the C-3 and C-4 of the galactose residue, but not the glucose residue, of lactose mainly interact with SSA (*29*).

Detailed view of the SSA–Neu5Acα2,6lactose complex in Fig. 5 indicated that the $O\gamma$ of S_{197} of SSA makes a hydrogen bond between the carboxyl group of sialic acid, as similar to the one between the serine residue (S_{114}) of WGA and sialic acid in the crystal structure of WGA–sialic acid complex (PDB ID: 1WGC). On the other hand, the corresponding amino-acid residue in ricin is glutamic acid (E_{199}) that seems to cause electrostatic repulsion between the negative charge of the carboxyl group in the sialic acid, supporting the inability of ricin to interact with sialylated oligosaccharides. Formation of another hydrogen bond was suggested between the $N\varepsilon$ of Q_{234} and the hydroxyl group attached to C-9 of the sialic acid. Again, the corresponding amino-acid residue in ricin is alanine (A_{237}) that cannot form any hydrogen bond, adding another support for the

Downloaded from http://jb.oxfordjournals.org/ at University of Science and Technology of China on September 28, 2012

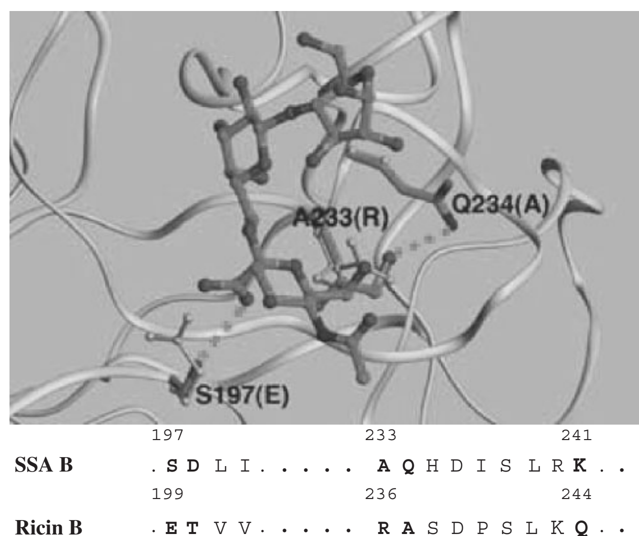

Fig. 5. **Model of SSA and Neu5Acα2,6lactose complex.** (A) Main chain of SSA is displayed by yellow tube model. Neu5Acα2,6lactose is shown by coloured ball-and-stick in each atom (carbon, green; nitrogen, blue; oxygen, red). Side chains of the amino-acid residues in SSA that are critical for the binding to sialic acid are indicated by coloured stick for each atom. Side chains of ricin corresponding to those of SSA are coloured by pink and their names are shown in parentheses. Cyan dotted lines indicate the hydrogen bonds. (B) Alignment of S_{197} to K_{241} of SSA B subunit to the consensus sequence of ricin B subunit. The bold character indicated the amino-acid residues subjected for the site-directed mutagenesis.

inability of ricin to bind sialic acid. While the conversion of R_{236} in ricin to A_{233} in SSA created a space to dock the glycerol group of sialic acid, the steric conflict between the glycerol group and the bulky R_{236} residue seems to contribute to the inability of ricin to interact with sialic acid.

To examine whether these amino-acid residues indicated by the docking simulation are really involved in the interaction with sialic acid, we replaced these amino acids with the corresponding ones in ricin, namely, S_{197} to E, A_{233} to R and Q_{234} to A, by introducing necessary mutation into the plasmid vector for SSA-BC expression. After the expression in *E. coli* and the following purification, the binding characteristics of the mutated SSA-BC proteins were analysed by using a biosensor based on surface plasmon resonance (Biacore).

As shown in Fig. 6 and Table 1, the expressed SSA-BC domain showed a high-affinity binding to fetuin, although the *K*a value was slightly lower than that of the native SSA. On the contrary, all the mutated proteins, S_{197}E, A_{233}R and Q_{234}A, showed very low binding to fetuin, which did not allow the kinetic treatment. These results clearly demonstrated, as expected from the docking simulation, that these amino-acid residues are critically important for the binding to sialic acid residues in the 2,6-linked sialylated oligosaccharides. In other words, the substitution of these amino-acid residues from those of ricin-like molecules to SSA seems to generate a lectin that specifically binds to Neu5Acα2,6Gal/GalNAc sequence.

From the docking simulation, several other amino-acid residues in the SSA-BC domain were also indicated for the possible interaction with the Neu5Ac of 2,6-sialyllactose. K_{241} and D_{198} in SSA, which correspond to Q_{244} and T_{200} in ricin, respectively, were indicated to interact with C-8 and NH group of Neu5Ac residue (Fig. 5). However, the replacement of these amino-acid residues to those corresponding amino acids in ricin did not affect the binding specificity of SSA-BC protein (Table 1).

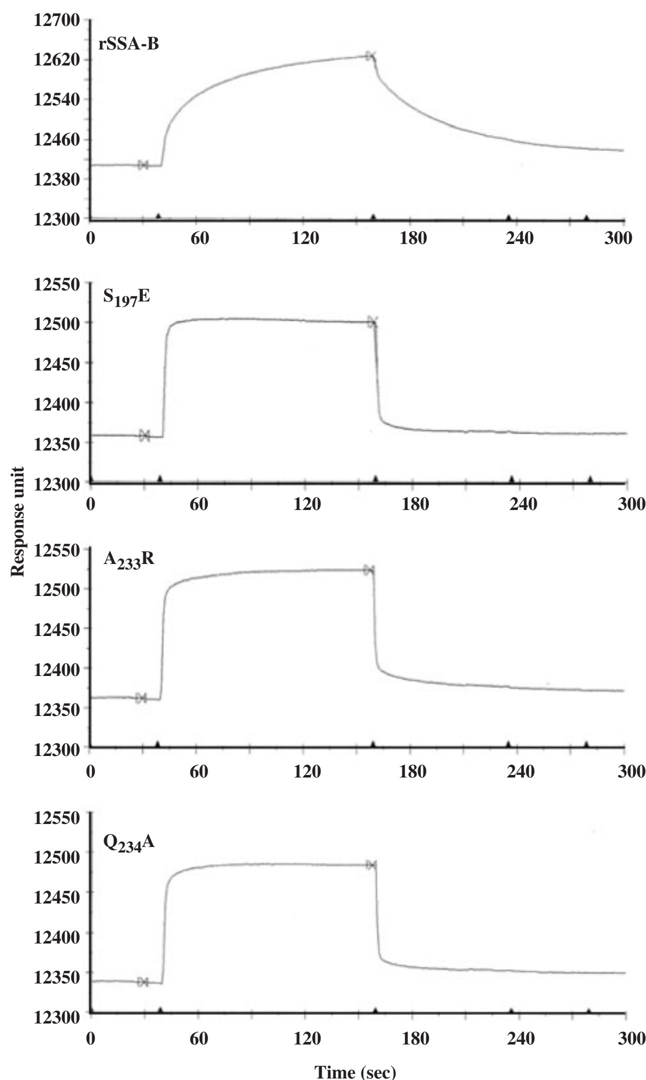

Fig. 6. **Interaction of mutated rSSA-BC proteins with fetuin.** Interaction of rSSA-BC and mutated proteins with fetuin was analysed by Biacore. These proteins (10 µg/ml) were applied to the sensor chip with the immobilized fetuin following the operation manual. The expressions such as S_{197}E, A_{233}R, Q_{234}A, indicate the mutation introduced to the corresponding proteins.

As the mutated proteins might lose the ability to bind to fetuin, we used the affinity chromatography on immobilized anti-SSA antibody, instead of the fetuin–agarose used for SSA/rSSA-BC purification, for the purification of the mutated proteins. Although we used different approaches for the purification of these proteins, the fact that the mutated proteins for K_{241} and D_{198} showed comparable affinity to the wild type

Downloaded from http://jb.oxfordjournals.org/ at University of Science and Technology of China on September 28, 2012

Table 1. **Kinetics data for the binding fetuin to SSA/rSSA-BC/the directed mutagenesis SSA-BC.**

	K_a ($M^{-1}S^{-1}$)[a]	K_d (S^{-1})[b]	K_a/K_d(M^{-1})[c]
SSA	2.78×10^4	3.96×10^{-3}	7.02×10^6
rSSA-BC	1.29×10^4	1.89×10^{-2}	6.83×10^5
$S_{197}E$	—	—	—
$A_{233}R$	—	—	—
$Q_{234}A$	—	—	—
$D_{198}T$	1.74×10^4	3.11×10^{-2}	5.59×10^5
$K_{241}Q$	1.63×10^4	4.00×10^{-2}	4.08×10^5

[a]Association rate constant. [b]Dissociation rate constant. [c]Association constant.

rSSA-BC indicated the difference of the purification method did not affect the binding specificity of these proteins.

DISCUSSION

In this article, we identified critical amino-acid residues required for the recognition of sialic acid by the Japanese elderberry bark lectin, SSA, by using molecular modelling/docking simulation and site-directed mutagenesis. As all the three amino-acid residues critical for the binding of SSA to sialic acid, S_{197}, A_{233} and Q_{234}, are conserved in the corresponding position of two other Neu5Acα2,6Gal/GalNAc-specific elderberry lectins, SNA I (*13*) and SNA I_f (*9*), the finding seems to be applicable for these lectins too. Replacement of one of these amino-acid residues to the one in the corresponding position of ricin B chain completely abolished the binding ability to fetuin. These results indicated that the mutation of the corresponding amino-acid residues in a putative ancestral lectin, which might be structurally related to ricin/abrin type lectins, to these amino-acid residues enabled SSA to generate a novel sialic acid binding site adjacent to the Gal/GalNAc binding site that is conserved in the ricin-like molecules (*15*).

The notion that SSA uses Gal/GalNAc binding site corresponding to the 2γ subdomain in the ricin B chain (*14*, *15*) for the binding to galactose residue in Neu5Acα2,6lactose was supported from the following observations. (i) SSA-BC domain that include the sequence corresponding to the 2γ subdomain of ricin carries the carbohydrate binding site; (ii) Critical amino-acid residues for the binding to Gal/GalNAc in the 2γ subdomain of ricin, such as D_{234}, Y_{248}, N_{252} and Q_{256}, were fully conserved in the corresponding positions of SSA-BC domain; (iii) Elderberry lectins such as SNA and SSA behave basically as Gal/GalNAc binding lectins (*2*, *3*) that corresponds to the specificity of 2γ subdomain of ricin but not to that of Gal-specific 1α subdomain of ricin (*15*); (iv) The indication of the amino-acid residues involved in the binding to sialic acid by the docking simulation based on such assumption was supported experimentally by the site-directed mutagenesis. It should be worth to show experimentally that the mutated SSA that lost the ability to sialic acid still retains the binding ability to Gal/GalNAc for further support of the hypothesis.

Concerning to the specific binding of the elderberry bark lectins to 2,6-linked sialylated oligosaccharides, but not to 2,3-linked sugars, the docking simulation in Fig. 4 clearly showed the difficulty to bind to 2,3-linked forms because of the steric hindrance between the sialic acid residues with the side chains of several amino-acid residues. Additionally, the substitution of the equatorial hydroxyl group at C-3 position of Gal/GalNAc residues in these oligosaccharides by sialic acid also seems to cause the inability to interact with these lectins. As discussed already, these elderberry lectins share a basic property as a Gal/GalNAc binding lectin that requires the interaction with the equatorial hydroxyl group at the C-3 position and the axial hydroxyl group at the C-4 position of Gal/GalNAc for the binding (*2*, *3*, *29*). Thus, the substitution of the hydroxyl group at the C-3 position should decrease the fitness of Gal/GalNAc residue to the binding site in the elderberry bark lectins.

Interestingly, the situation seems to be different in the case of 2,3-linked sialylated oligosaccharide-specific lectins from *Maackia amurensis* seeds, MAL and MAH (*10*, *11*, *30*). In the case of *Maackia amurensis* lectins, substitution of the equatorial hydroxyl group at the C-3 position of galactose did not diminish the binding, rather enhanced the affinity of the corresponding sugars to the lectins. Yamamoto *et al*. (*31*) discussed the differences between the elderberry bark lectins and *Maackia amurensis* lectins from the differences in the dominant ligand for these lectins, that is, Gal/GalNAc is the dominant sugar for the elderberry bark lectins, whereas sialic acid is the dominant sugar for the *Maackia amurensis* lectins.

What could be the outcome of such development of carbohydrate binding specificity in the elderberry bark lectins for plant itself? Plant lectins and RIPs have been considered as a part of defense machinery in plants (*32*, *33*). In the case of traditional type 2 RIPs, the carbohydrate binding subunit (B chain) serves to deliver the toxic subunit (A chain) to the target cells by binding to the cell surface carbohydrates with terminal galactose residues (*34*). If the elderberry bark lectins also bind to the target cells and are internalized as similar to other type 2 RIPs, it would be intriguing to consider that the binding of the B chain to sialoglyco-conjugates favours to deliver the toxic subunit to mammalian cells specifically, because such molecules are abundant on the mammalian cell surface but not present in the plant cells. However, such a speculation has a major drawback because the A chain of the elderberry bark lectins showed almost no ribosome inactivating activity, though structurally related to the A chains of traditional type 2 RIPs (*12*, *13*).

In addition to the Neu5Acα2,6Gal/GalNAc-specific bark lectins discussed here, the elderberry plants belonging to the genus *Sambucus* produces various types of lectins and RIP-related proteins. They produce a truncated polypeptide corresponding to a part of the Neu5Aα2,6Gal/GalNAc-binding subunit (*9*), Gal/GalNAc-specific lectins (*35*, *36*), type 2 RIPs with *in vitro* ribosome-inactivating activity but not toxicity (*37–39*), a type 2 RIP with inactive B chain (*40*) and a monomeric lectin corresponding to the B chain of typical type 2 RIPs (*26*). Similarity in the sequence of those genes encoding these elderberry proteins clearly indicates that they evolved from a small number of ancestral genes. Although it has been postulated that the presence of

Downloaded from http://jb.oxfordjournals.org/ at University of Science and Technology of China on September 28, 2012

a wide repertoire of these lectins/RIPs favours the plants to protect themselves from various types of predators, validity of such a speculation should further be examined in future studies.

The finding of amino-acid residues critical for the interaction between elderberry lectins and sialylated oligosaccharides reported here should contribute to understand how these lectins obtained such a unique binding specificity during molecular evolution and also to design novel lectins/RIPs that bind to sialoglycoconjugates specifically. Very recently, Yabe *et al*. (*41*) reported the creation of a sialic acid binding lectin by error-prone PCR based on the assumption that various naturally occurring sialic acid binding lectins have evolved from the ancestral Gal-binding lectins. The findings reported in this paper coincide with this assumption and should help to develop such an approach to create novel and useful lectins.

We thank Dr Julie E. Krawetz for her help in the construction of the plasmids for site-directed mutagenesis. This work was partly supported by a MEXT-HAITEKU grant from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

REFERENCES

1. Rudiger, H. and Gabius, H.J. (2001) Plant lectins: occurrence, biochemistry, functions and applications. *Glycoconj. J.* **18**, 589–613
2. 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
3. Shibuya, N., Tazaki, K., Song, Z.W., Tarr, G.E., Goldstein, I.J., and Peumans, W.J. (1989) A comparative study of bark lectins from three elderberry (*Sambucus*) species. *J. Biochem.* **106**, 1098–1103
4. Shibuya, N., Goldstein, I.J., Broekaert, W.F., Nsimba-Lubaki, M., Peeters, B., and Peumans, W.J. (1987) Fractionation of sialylated oligosaccharides, glycopeptides, and glycoproteins on immobilized elderberry (*Sambucus nigra* L.) bark lectin. *Arch. Biochem. Biophys.* **254**, 1–8
5. Kaku, H. and Shibuya, N. (1992) Preparation of a stable subunit of Japanese elderberry (*Sambucus sieboldiana*) bark lectin and its application for the study of cell surface carbohydrates by flow cytometry. *FEBS Lett.* **306**, 176–180
6. Nakajima, K., Kinoshita, M., Oda, Y., Masuko, T., Kaku, H., Shibuya, N., and Kakehi, K. (2004) Screening method of carbohydrate-binding proteins in biological sources by capillary affinity electrophoresis and its application to determination of *Tulipa gesneriana* agglutinin in tulip bulbs. *Glycobiology* **14**, 793–804
7. Sata, T., Lackie, P.M., Taatjes, D.J., Peumans, W., and Roth, J. (1989) Detection of the Neu5 Ac (alpha 2,3) Gal (beta 1,4) GlcNAc sequence with the leukoagglutinin from *Maackia amurensis*: light and electron microscopic demonstration of differential tissue expression of terminal sialic acid in alpha 2,3- and alpha 2,6-linkage. *J. Histochem. Cytochem.* **37**, 1577–1588
8. Schauer, R. (2000) Achievements and challenges of sialic acid research. *Glycoconj. J.* **17**, 485–499
9. Peumans, W.J., Roy, S., Barre, A., Rouge, P., van Leuven, F., and van Damme, E.J. (1998) Elderberry (*Sambucus nigra*) contains truncated Neu5Ac(alpha-2,6)Gal/GalNAc-binding type 2 ribosome-inactivating proteins. *FEBS Lett.* **425**, 35–39
10. Knibbs, R.N., Goldstein, I.J., Ratcliffe, R.M., and Shibuya, N. (1991) Characterization of the carbohydrate binding specificity of the leukoagglutinating lectin from *Maackia amurensis*. Comparison with other sialic acid-specific lectins. *J. Biol. Chem.* **266**, 83–88
11. Wang, W.C. and Cummings, R.D. (1988) The immobilized leukoagglutinin from the seeds of *Maackia amurensis* binds with high affinity to complex-type Asn-linked oligosaccharides containing terminal sialic acid-linked alpha-2,3 to penultimate galactose residues. *J. Biol. Chem.* **263**, 4576–4585
12. 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
13. Van Damme, E.J., Barre, A., Rouge, P., Van Leuven, F., and Peumans, W.J. (1996) The NeuAc(alpha-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
14. Rutenber, E. and Robertus, J.D. (1991) Structure of ricin B-chain at 2.5 A resolution. *Proteins* **10**, 260–269
15. Sphyris, N., Lord, J.M., Wales, R., and Roberts, L.M. (1995) Mutational analysis of the Ricinus lectin B-chains. Galactose-binding ability of the 2 gamma subdomain of *Ricinus communis* agglutinin B-chain. *J. Biol. Chem.* **270**, 20292–20297
16. Chen, Y., Rouge, P., Peumans, W.J., and van Damme, E.J. (2002) Mutational analysis of the carbohydrate-binding activity of the NeuAc(alpha-2,6)Gal/GalNAc-specific type 2 ribosome-inactivating protein from elderberry (*Sambucus nigra*) fruits. *Biochem. J.* **364**, 587–592
17. Tazaki, K. and Shibuya, N. (1989) Purification and partial characterization of a lectin from the bark of Japanese elderberry (*Sambucus sieboldiana*). *Plant Cell Physiol.* **30**, 899–903
18. Kanaoka, M., Kishimoto, F., Ueki, Y., and Umeyama, H. (1989) Alignment of protein sequences using the hydrophobic core scores. *Protein Eng.* **2**, 347–351
19. Needleman, S.B. and Wunsch, C.D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443–453
20. Umezawa, Y. and Umeyama, H. (1988) Computer screening and visualization of hydrophobic core of protein. *Chem. Pharm. Bull.* **36**, 4652–4658
21. Ponder, J.W. and Richards, F.M. (1987) Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* **193**, 775–791
22. Richardson, P.T., Roberts, L.M., Gould, J.H., and Lord, J.M. (1988) The expression of functional ricin B-chain in *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* **950**, 385–394
23. de Sousa, M., Roberts, L.M., and Lord, J.M. (1999) Restoration of lectin activity to an inactive abrin B chain by substitution and mutation of the 2 gamma subdomain. *Eur. J. Biochem.* **260**, 355–361
24. Vitetta, E.S. and Yen, N. (1990) Expression and functional properties of genetically engineered ricin B chain lacking galactose-binding activity. *Biochim. Biophys. Acta* **1049**, 151–157
25. Chang, M.S., Russell, D.W., Uhr, J.W., and Vitetta, E.S. (1987) Cloning and expression of recombinant, functional ricin B chain. *Proc. Natl. Acad. Sci. USA* **84**, 5640–5644
26. Rojo, M.A., Kaku, H., Ishii-Minami, N., Minami, E., Yato, M., Hisajima, S., Yamaguchi, T., and Shibuya, N. (2004) Characterization and cDNA cloning of monomeric lectins that correspond to the B-chain of a type 2 ribosome-inactivating protein from the bark of Japanese elderberry (*Sambucus sieboldiana*). *J. Biochem.* **135**, 509–516

Downloaded from http://jb.oxfordjournals.org/ at University of Science and Technology of China on September 28, 2012

27. Wright, C.S. (1992) Crystal structure of a wheat germ agglutinin/glycophorin-sialoglycopeptide receptor complex. Structural basis for cooperative lectin-cell binding. *J. Biol. Chem.* **267**, 14345–14352
28. Wright, C.S. (1990) 2.2 A resolution structure analysis of two refined N-acetylneuraminyl-lactose–wheat germ agglutinin isolectin complexes. *J. Mol. Biol.* **215**, 635–651
29. Takesada, H., Shibuya, N., and Nagashima, N. (1992) The interaction of elderberry (*Sambucus sieboldiana*) bark lectin and sialyloligosaccharides as detected by 1H-NMR. *J. Biochem.* **112**, 143–146
30. Kawaguchi, T., Matsumoto, I., and Osawa, T. (1974) Studies on hemagglutinins from *Maackia amurensis* seeds. *J. Biol. Chem.* **249**, 2786–2792
31. Yamamoto, K., Konami, Y., and Irimura, T. (1997) Sialic acid-binding motif of *Maackia amurensis* lectins. *J. Biochem. (Tokyo)* **121**, 756–761
32. Peumans, W.J. and Van Damme, E.J. (1995) Lectins as plant defense proteins. *Plant Physiol.* **109**, 347–352
33. Peumans, W.J., Hao, Q., and Van Damme, E.J. (2001) Ribosome-inactivating proteins from plants: more than RNA N-glycosidases? *FASEB J.* **15**, 1493–1506
34. Barbieri, L., Battelli, M.G., and Stirpe, F. (1993) Ribosome-inactivating proteins from plants. *Biochim. Biophys. Acta* **1154**, 237–282
35. Kaku, H., Peumans, W.J., and Goldstein, I.J. (1990) Isolation and characterization of a second lectin (SNA-II) present in elderberry (*Sambucus nigra* L.) bark. *Arch. Biochem. Biophys.* **277**, 255–262
36. Rojo, M.A., Citores, L., Jimenez, P., Ferreras, J.M., Arias, F.J., Mendez, E., and Girbes, T. (2003) Isolation and characterization of a new d-galactose-binding lectin from *Sambucus racemosa* L. *Protein Pept Lett* **10**, 287–293
37. Girbes, T., Citores, L., Iglesias, R., Ferreras, J.M., Munoz, R., Rojo, M.A., Arias, F.J., Garcia, J.R., Mendez, E., and Calonge, M. (1993) Ebulin 1, a nontoxic novel type 2 ribosome-inactivating protein from *Sambucus ebulus* L. leaves. *J. Biol. Chem.* **268**, 18195–18199
38. Rojo, M.A., Yato, M., Ishii-Minami, N., Minami, E., Kaku, H., Citores, L., Girbes, T., and Shibuya, N. (1997) Isolation, cDNA cloning, biological properties, and carbohydrate binding specificity of sieboldin-b, a type II ribosome-inactivating protein from the bark of Japanese elderberry (*Sambucus sieboldiana*). *Arch. Biochem. Biophys.* **340**, 185–194
39. Citores, L., De Benito, F.M., Iglesias, R., Ferreras, J.M., Argueso, P., Jimenez, P., Testera, A., Camafeita, E., Mendez, E., and Girbes, T. (1997) Characterization of a new non-toxic two-chain ribosome-inactivating protein and a structurally-related lectin from rhizomes of dwarf elder (*Sambucus ebulus* L.). *Cell Mol. Biol. (Noisy-le-grand)* **43**, 485–499
40. Van Damme, E.J., Barre, A., Rouge, P., Van Leuven, F., and Peumans, W.J. (1997) Isolation and molecular cloning of a novel type 2 ribosome-inactivating protein with an inactive B chain from elderberry (*Sambucus nigra*) bark. *J. Biol. Chem.* **272**, 8353–8360
41. 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

Downloaded from http://jb.oxfordjournals.org/ at University of Science and Technology of China on September 28, 2012