Diverse Sugar-Binding Specificities of Marine Invertebrate C-Type Lectins

Hiroki Matsubara, Sachiko Nakamura-Tsuruta, Jun Hiranayashi, Mitsuru Jimbo, Hisao Kamiya, Tomohisa Ogawa, and Koji Muramoto

1Department of Biomolecular Science, Graduate School of Life Sciences, Tohoku University, Sendai 981-8555, Japan
2Research Center for Glycoscience, National Institute of Advanced Industrial Science and Technology, Tsukuba 305-8568, Japan
3School of Fisheries Sciences, Kitasato University, Ohfunato 022-0101, Japan

Received October 3, 2006; Accepted November 12, 2006; Online Publication, February 7, 2007
[doi:10.1271/bbb.60543]

The sugar-binding specificities of C-type lectins isolated from marine invertebrates were investigated by frontal affinity chromatography (FAC) using 100 oligosaccharides. The lectins included BRA-2 and BRA-3, multiple lectins from the hemolymph of the acorn barnacle, Megabalanus rosa, and BRL from the acorn barnacle, Balanus rostatus. The diverse sugar-binding specificities of the C-type lectins were determined by FAC analysis. BRA-2 recognized α2-6 sialylation but not α2-3 sialylation on glycans. On the other hand, BRA-3 showed high affinity for oligosaccharides with α-linked non-reducing terminal galactose, but not for sialylated forms, and BRL showed enhanced recognition activity towards Lewis^a and Lewis^x epitopes.

Key words: carbohydrate binding specificity; C-type lectin; frontal affinity chromatography; invertebrate lectin; lectin

Lectins are a group of sugar-binding proteins that recognize specific carbohydrate structures and agglutinate various cells by binding to cell-surface glycoconjugates. Animals produce a variety of lectins, many of which have been implicated in cell recognition phenomena. Most animal lectins are classified into several lectin families, depending on their sequence similarities and sugar-binding specificities: C-type lectins, galectins, etc. The C-Type lectin family was named for its Ca^{2+}-dependent sugar-binding activity, and biochemical studies revealed that the lectin family shares a compact globular module structure, the carbohydrate-recognition domain (CRD). This domain is called the C-type CRD or C-type like domain (CTLD), and it contains a characteristic double-loop (loop in a loop) stabilized by two highly conserved disulfide bonds. As the number of determined amino acid sequences grew, it became clear that the superfamily of proteins containing CTLDs is a large group of extracellular Metazoan proteins with diverse functions. The family members found in vertebrates have been classified into over 14 groups based on their overall domain architectures: e.g., lecticans, the asialoglycoprotein receptor group, collectins, selectins, natural killer (NK)-cell receptors, etc. Each family member shows a unique multi-domain structure and ligand-binding activity.

In contrast to vertebrate C-type lectins, those in invertebrates are generally composed only of CTLD, and are assumed to participate in the host defense system in place of immunoglobulins in vertebrates. Recent progress in the genome-sequencing project for the invertebrates Drosophila melanogaster and Caenorhabditis elegans indicates that the CTLD repertoires of these species are drastically different from each other, and from the known vertebrate groups; the superfamily has undergone very few changes in the 450 million years of vertebrate radiation. Various C-type lectins have been isolated and characterized from many marine invertebrates, including barnacles, a sea urchin, sea cucumbers, a starfish, a sponge, tunicates, etc. It is generally accepted that they are specific for ϕ-galactose or N-acetyl-ϕ-galactosamine, as determined by inhibition assay of hemagglutination using simple sugars. There are few invertebrate lectins, of which the detailed sugar-binding specificities and endogenous ligands are known, though it is necessary to elucidate them to understand their biological functions.

We have proposed possible multiple functions of C-type lectins in acorn barnacle, Megabalanus rosa, hemolymph: their participation in defense and biomineralization. The lectins comprise three different
molecular forms with respect to molecular weight, designated BRA-1, BRA-2, and BRA-3.\textsuperscript{18} On the other hand, only one C-type lectin, designated BRL, has been isolated from the hemolymph of the acorn barnacle \textit{Balanus rostratus}.\textsuperscript{7} The C-type lectins from acorn barnacles agglutinate various animal blood cells and tumor cells. The binding of the lectins to rabbit erythrocyte ghosts is inhibited by various simple sugars, including D-galactose, D-galactosamine, N-acetyl-D-galactosamine, and N-acetylneuraminic acid, indicating that they have wide sugar-binding specificities and can function as recognition factors for a variety of foreign substances.\textsuperscript{18}

Inhibition assay of hemagglutination with simple sugars has been used to examine the sugar-binding specificities of most lectins, because no simple method is available to do this with a large repertoire of oligosaccharides. To investigate the detailed sugar-binding specificities of lectins, the frontal affinity chromatography (FAC) system was recently reinforced by an efficient combination of the advantages of high-performance liquid chromatography (HPLC) and a large repertoire of pyridylaminated (PA)-oligosaccharides.\textsuperscript{19,20} In this study, we applied the automated FAC system to explore the detailed sugar-binding specificities of C-type lectins isolated from acorn barnacles. The FAC system indicated that the marine invertebrate C-type lectins BRAs and BRL had diverse binding specificities toward oligosaccharides.

**Materials and Methods**

**Materials.** Acorn barnacle lectins, BRA-1, BRA-2, BRA-3, and BRL, were purified from the hemolymph of \textit{M. rosa} and \textit{B. rostratus} as previously described.\textsuperscript{21}

Pyridylaminated (PA-) and \textit{p}-nitrophenyl (pNP-) oligosaccharides were obtained commercially from Takara Bio (Kyoto, Japan), Seikagaku Kogyo (Tokyo, Japan), and TRC (North York, Canada), and their structures were numbered as described (Fig. 1).\textsuperscript{22} PA-911 (LacNAc) was pyridylaminated with GlycoTag (Takara Bio) before use. pNP-6-’Sialyllactose (6’SiaLac, Neu5Aca2-6Galβ1-4Glc-) was the generous gift of Dr. Usui and Dr. Murata (Shizuoka University, Shizuoka, Japan).

**Preparation of lectin columns.** BRA-1, BRA-2, BRA-3, and BRL individually were dissolved in 20 mM HEPES–NaOH (pH 7.5) containing 0.15 M NaCl and 10 mM CaCl\(_2\), and coupled to NHS-activated Sepharose 4FF (Amersham Pharmacia Biotech) according to the manufacturer’s instructions. Lectin-Sepharose was suspended in 10 mM Tris–HCl (pH 7.4) containing 0.8% NaCl and 10 mM CaCl\(_2\) (running buffer). The slurry was packed into a capsule-type miniature column (ϕ 2 × 10 mm, 31 μl) specially designed for FAC.\textsuperscript{22}

**Frontal affinity chromatography.** FAC was performed using an automated system (FAC-1), as described previously.\textsuperscript{22} Lectin columns were equilibrated with the running buffer before FAC analysis, which was performed using PA-oligosaccharides (2.5 nM) or pNP-oligosaccharides (5 μM) dissolved in the running buffer. The flow rate and the column temperature were kept at 0.125 ml/min and 25°C. An excess volume (0.5–0.8 ml) of each oligosaccharide solution was successively injected into the column by an auto-sampling system. Elution of pNP- and PA-oligosaccharide was monitored by measuring UV (280 nm) and fluorescence (ex/em = 310/380 nm) respectively.

Concentration-dependence analysis and a subsequent Woolf-Hofstee-type plot, (V–V\(_0\)) vs. (V–V\(_0\))[A]\(_0\), were determined to 
\[B(t) = \frac{V(t)}{V_0(t)} = \frac{[A]_0}{K_d[V_0] - [A]_0}\]

**Results**

**Evaluation of lectin columns.** The lectins were immobilized onto NHS-activated Sepharose 4FF in 20 mM HEPES–NaOH (pH 7.5) containing 0.15 M NaCl in the presence of 10 mM CaCl\(_2\). The presence of Ca\(^{2+}\) in the coupling buffer was essential to stabilize the lectins during the coupling reaction. The amounts of immobilized BRAs and BRL onto gels were estimated to be about 5 mg/ml gel. From Woolf-Hofstee-type plots, the \(B_t\) and \(K_d\) values were determined to be 0.93 nmol and 0.6 × 10\(^{-4}\) M for the BRA-2 column, 1.37 nmol and 0.46 × 10\(^{-4}\) M for the BRA-3 column, and 4.09 nmol and 2.4 × 10\(^{-4}\) M for the BRL column, respectively. The availabilities of the BRA-2, BRA-3, and BRL columns were calculated to be 13, 14, and 52%, respectively. Availability was calculated by the following formula: 
\[\text{Availability} = \frac{B_t \times \text{immobilized lectin (g)}}{\text{molecular weight (g/mol)} \times 100\%}\]

Availability was improved about 2-fold by changing the coupling buffer from 0.2 M NaHCO\(_3\) (pH 8.3) containing 0.5 M NaCl to HEPES (pH 7.5) containing 0.15 M NaCl and 0.01 M CaCl\(_2\). Although the BRA-1 column showed a binding profile similar to that of the BRA-2 column, the \(B_t\) and \(K_d\) values of the BRA-1 column could not be determined, because none of the pNP-saccharides tested showed significant affinity as required for concentration-dependence analysis.
A.  N-Linked glycans

B. Glycolipid-type glycans

C. Others

Fig. 1. Schematic Representation of PA-Oligosaccharides Used in This Study.
The reducing terminals are pyridylaminated. Symbols used to represent pyranose rings of monosaccharides are shown in the box at the bottom of the figure.
should be noted that BRA-2 had high affinity to some Gal, such as Gb3 (715) and Galili penta (725), and the small Arabic figures correspond to the sugar numbers indicated in Fig. 1. Large Roman figures at the bottom represent types of glycans: N-linked glycans from 002 to 017, high-mannose-type (group I), from 101 to 203, agalacto-type (group II), from 301 to 420, galactosylated-type (group III), and from 501 to 602, sialylated-type (group IV). N-linked glycans: from 701 to 739, glycolipid-type glycans (group V), and from 901 to 911, others (group VI).

Common feature of sugar-binding specificity of barnacle lectins

The $K_d$ values of lectins for 100 PA-glycans, including 52 N-linked glycans, 38 glycolipid-type glycans, and 10 other-type glycans, are summarized in Figs. 2–4. For better comparison of affinity strength, a bar graph was constructed using $K_d$ ($= 1/K_a$) values, though $K_a$ is used in the text. All barnacle lectins showed high affinity against bi-antennary galactosylated-type oligosaccharides (307, 313, 314, 323, 405, 410, 418, 419, and 420) and glycolipid-type oligosaccharides having Gal$\alpha$1-3 or 4 Gal, such as Gb3 (715) and Galili penta (725), and those having Gal$\beta$1-3 Gal, such as $\beta$GalLac (735) (Fig. 2). Lacto-N-hexaose (733, LNH) and lacto-N-neohexaose (734, LNeH) were also recognized by the lectins; i.e., the lectins had binding affinity for galactosylated-type $N$-glycans, but not at all for high mannose-type and agalacto-type $N$-glycans.

**BRA-2**

BRA-2 showed extensive affinity for galactosylated-type $N$-glycans. The affinity increased with increases in the branching number up to tetra-antennary $N$-glycans, viz., mono-antennary (301, $K_d = 117 \mu M$), biantennary (307, 56 $\mu M$), triantennary (313, 26 $\mu M$), and tetra-antennary (323, 17 $\mu M$). $\alpha$-1,6Fucosylation of the $N$-glycans did not affect the interaction with BRA-2. It should be noted that BRA-2 had high affinity to some sialylated glycans: mono-sialylated bi-antennary glycans (501, 502), di-sialylated bi-antennary glycans (503), and tri-sialylated tri-antennary glycans (504) (Fig. 3). $\alpha$2-6Sialylation of bi-antennary glycans (307) enhanced affinity for BRA-2. Affinity increased with the number of $\alpha$2-6 linked 5NeuAc up to tri-sialylated tri-antennary: $\beta$GalLac (735), mono-sialylated glycans (501, $K_d = 27 \mu M$), mono-sialylated glycans (502, 28 $\mu M$), and di-sialylated glycans (503, 16 $\mu M$). 6'SialLac (704, 123 $\mu M$) in glycolipid-type glycans also showed affinity for BRA-2. The highest affinity was observed for 504 ($K_d = 7 \mu M$). On the other hand, $\alpha$2-3 sialylation of glycans reduced the affinity of BRA-2; $\alpha$2-3 mono-sialylated glycans (601, 169 $\mu M$) and $\alpha$2-3 di-sialylated glycans (602, no affinity). This result indicates that BRA-2 discriminates $\alpha$2-6 sialylated glycans from $\alpha$2-3 sialylated glycans.

**BRA-3**

BRA-3 gave a binding affinity profile different to that of BRA-2. The former showed no affinity for mono-galactosylated bi-antennary (301–305) or sialylated-type (501–504) glycans (Fig. 2). $\alpha$1-6 (core) fucosylation of the N-glycans slightly changed the binding affinity of BRA-3: $\beta$GalLac (307, $K_d = 250 \mu M$) versus 405 (158 $\mu M$) and 323 (36 $\mu M$) versus 418 (42 $\mu M$).

BRA-3 showed the highest affinity for T-antigen (911, 25 $\mu M$) (Fig. 4). The incorporation of $\alpha$-linked non-reducing terminal galactose to lactose (701, 256 $\mu M$)

![Bar Graph Representation of Association Constants $K_d (= 1/K_a)$ of BRA-2 (top), BRA-3 (middle), and BRL (bottom) toward N-Linked Glycans.](image)
significantly increased the binding affinity of BRA-3, as found with Gb3 (715, 33 μM). Fucosylation and sialylation of lactose (701) reduced the affinity of BRA-3: 704 (no affinity), 705 (558 μM), and 718 (no affinity).

**BRL**

BRL increased the binding affinity with the branching number up to tetra-antennary galactosylated-type glycans: viz., 301 ($K_d = 566$ μM), 307 (189 μM), 313 (90 μM), and 323 (33 μM), whereas the affinity decreased with the α2-6 sialylation of bi-antennary glycan (307): viz., 501 ($K_d = 247$ μM), 502 (302 μM), and 503 (no affinity). The α1-6 (core) fucosylation of mono-antennary and tetra-antennary glycan affected affinity to some extent; *i.e.* 301 (566 μM) versus 401 (787 μM) and 313 (90 μM) versus 410 (54 μM) respectively. The α1-3 fucosylation of lacto-N-neotetraose (724, 274 μM) also affected affinity: lacto-N-fucosyl pentaose (726, 178 μM) (Fig. 4). In contrast to BRA-2 and BRA-3, BRL recognized Lewis$^a$ (726) and Lewis$^b$ (730) epitopes.

**Discussion**

Vertebrate C-type lectins are generally composed of multi-functional domain structures to express a variety of biological functions in addition to CTLD.
carbohydrate binding sites of C-type lectins used to be classified as Gal-type or Man-type according to the presence of the sequence motif Gln-Pro-Asp (QPD) or Glu-Pro-Asn (EPN) respectively. C-Type lectins in marine invertebrates are composed of simple domain structures; most of them are composed only of CTLD. The invertebrate lectins show low sequence homologies, of less than 40%, to one another even among closely related species, and have sugar-binding specificity toward D-galactose or N-acetyl-D-galactosamine regardless of the presence of the sequence motif. For example, perlucin from the abalone Haliotis laevigata, containing the QPD motif, recognizes both D-galactose and D-mannose. The invertebrate lectins show low sequence homologies, of less than 40%, to one another even among closely related species, and have sugar-binding specificity toward D-galactose or N-acetyl-D-galactosamine regardless of the presence of the sequence motif, indicating that not all factors responsible for carbohydrate recognition have been identified, and that it is therefore difficult to predict binding selectivity by sequence comparisons and homology modeling.

This study demonstrated the diverse sugar-binding specificities of invertebrate C-type lectins using an automated FAC system. The multiple C-type lectins, BRA-1 (Mr 330k), BRA-2 (Mr 140k), and BRA-3 (Mr 64k), are major hemolymph proteins in M. rosa. BRA-1 and BRA-2 are composed of an identical subunit (22kDa), which consists of 173 amino acid residues with one N-glycosylation site at Asn-39. BRA-3 is composed of four identical subunits (16kDa), which consist of 138 amino acid residues. The amino acid sequences of BRA-2 and BRA-3 show only 16% identity. BRL is the only C-type lectin from the hemolymph of the acorn barnacle B. rostratus. BRL (Mr 120k) is a multimeric protein whose subunit (35 kDa) consists of 182 amino acid residues. BRL is 46 and 15% identical to BRA-2 and BRA-3 respectively. BRA-2 and BRL contain the QPD motif (residues 136–138 and 145–147 respectively), whereas BRA-3 presents the NPN sequence (residues 92–94). It is noteworthy that BRA-2 distinguishes α2-6Neu5Ac from α2-3Neu5Ac. Among the superfamily of proteins containing CTLDs, selectins are unique due to their affinity for sialyl-Lewisα (sLeα). However, compared to selectins, BRA-2 did not bind to sLeα, which contains α2-3 linked sialic acid. Moreover, BRA-2 did not recognize glycolipid type oligosaccharides, though BRA-3 and BRL widely recognized them with different properties. The sugar-binding specificity of BRA-1 was very similar to that of BRA-2, except for the affinity toward pNP-saccarides, suggesting that subunit structures affected the specificity (data not shown).

In conclusion, the present success of the FAC analysis of invertebrate C-type lectins was due to the preparation of the lectin columns, stabilizing the lectins by the addition of calcium ions to the coupling buffer. FAC analysis indicated the diverse sugar-binding specificities of C-type lectins from marine invertebrates. CTLDs, which are stabilized by two highly conserved disulfide bonds and have diverse sugar-binding specificities, might be an interesting and promising target for protein engineering of carbohydrate recognition modules.

Acknowledgments

We thank Dr. T. Usui and Dr. T. Murata (Faculty of Agriculture, Shizuoka University) for providing pNP-6’SiaLac. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (no. 17380122 to K.M.), by NEDO (the New Energy and Industrial Technology Organization) under METI (the Ministry of Economy, Trade, and Industry of Japan), and by a grant from The Salt Science Foundation (no. 0629 to K.M.).

References