J. Biochem. 121, 1002-1009 (1997)

The Two Lectin Domains of the Tandem-Repeat 32-kDa Galectin of the Nematode *Caenorhabditis elegans* Have Different Binding Properties. Studies with Recombinant Protein[1]

Yoichiro Arata,[2] Jun Hirabayashi, and Ken-ichi Kasai

Department of Biological Chemistry, Faculty of Pharmaceutical Sciences, Teikyo University, Sagamiko, Kanagawa 199-01

Received for publication, October 9, 1996

Some properties of recombinant proteins derived from the 32-kDa galectin isolated from the nematode *Caenorhabditis elegans*, which lectin is composed of two tandemly repeated homologous domains [Hirabayashi *et al.* (1992) *J. Biol. Chem.* 267, 15485], were studied in order to elucidate the function of this unique polypeptide architecture. We expressed the whole molecule (N32), the N-terminal lectin domain (Nh), and the C-terminal lectin domain (Ch) in *Escherichia coli* using the expression vector pET21a. All of the recombinant proteins were bound by asialofetuin-Sepharose. CD spectra of the recombinant proteins indicated all of them to be rich in β-structure and properly refolded. Gel filtration on an HPLC column suggested that all of them existed as monomers. Neither Nh nor Ch seemed to form dimers, in contrast to vertebrate proto-type galectins. Only N32 showed hemagglutination activity towards trypsinized rabbit erythrocytes. Comparison of the affinity of N32, Nh, and Ch for asialofetuin-Sepharose by frontal affinity chromatography [Kasai *et al.* (1986) *J. Chromatogr.* 376, 33] showed that Ch has 7-fold weaker affinity than N32, and Nh proved to have still weaker affinity. Since the Asn residue in the CRD (carbohydrate recognition domain), which is conserved in all other galectins, is substituted by Ser in the case of Nh, these data suggest that the two CRDs in this tandem-repeat galectin have different sugar binding properties and that the 32-kDa galectin may serve as a heterobifunctional crosslinker.

Key words: *C. elegans*, frontal affinity chromatography, β-galactoside-binding lectin, galectin, tandem-repeat-type galectin.

Most animal lectins can be divided into two major groups from the standpoint of primary structure (*1, 2*). One is the family of C-type (Ca^{2+}-dependent) lectins and the other is the family of galectins. Galectins are metal-independent, β-galactoside-binding lectins (*3-5*). The complete sequences of polypeptides (*6*) and cDNA (*7*) of these lectins were first determined for one of the chicken galectins. Though they used to be called S-type lectins, S-Lac lectins, galaptins, and so on, recently the term galectin was adopted as the general name (*8*). To be regarded as a member of the galectin family, a protein must fulfill two criteria: affinity for β-galactosides and a conserved sequence motif in the carbohydrate-binding site. A number is assigned to each mammalian galectin in the order of discovery. The complete primary structure has been reported so far for galectin-1 from humans (*9, 10*), rat (*11*), mouse (*12*), and cow (*13*), for galectin-2 from humans (*14*), for galectin-3 from humans (*15-17*), rat (*18*), and mouse (*19*), and for galectin-4 from rat (*20*) and pig (*21*). Recently galectin-5 from rat (*22*), galectin-7 from humans (*23, 24*), and galectin-8 from rat (*25*) have joined the family. Note that non-mammalian galectins are not included in this numbering system because of the difficulty of inferring the relationship between particular mammalian galectins and those found in non-mammalian species. As non-mammalian galectins, those from chicken (*6, 7, 26*), electric eel (*27*), and frogs [*Xenopus laevis* (*28*) and *Rana catesbeiana* (*29*)] have also been studied.

Though galectins had been reported only in vertebrates for a long time, we found that the primary structure of a 32-kDa lectin isolated from the nematode *Caenorhabditis elegans* is homologous to that of vertebrate galectins (*30, 31*). Since vertebrates belong to the deuterostomes, and nematodes to the proteostomes, this finding indicates that there should have been a common ancestor protein at the time of divergence of these two large branches of animals; that is, more than 800 million years ago. Therefore, galectins may have been playing some fundamental roles in the animal kingdom since the precambrian era, conserving their primary structures and sugar-binding function. More recently, the marine sponge *Geodia cydonium* (*32*) was reported to be the second invertebrate demonstrated to have galectins. Therefore, galectins are now known to be distributed from the simplest to the most complicated multicellular animals.

[1] This study was supported in part by Grants-in-Aid for Scientific Research (05274101, 07857168, 08772109) from the Ministry of Education, Science, Sports and Culture of Japan and by grants from the Mitsubishi Foundation, Naito Foundation, and Kihara Memorial Yokohama Foundation.

[2] To whom correspondence should be addressed. Tel: +81-426-85-3741, Fax: +81-426-85-3742

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

C. elegans is a very useful experimental animal for the study of the molecular events underlying differentiation and development. Extensive genetic analyses have been performed and a variety of mutants obtained (*33, 34*). It should, therefore, be an excellent system for the elucidation of the biological significance of galectin.

One of the remarkable characteristics of the *C. elegans* 32-kDa galectin is its molecular architecture. This lectin is composed of two domains, each of which is homologous to typical vertebrate 14-kDa-type galectins (about 25% homologies). It was the first discovery of not only an invertebrate galectin, but also of the existence of "tandem-repeat" type galectins. Consequently, galectins which have been characterized so far have turned out to be classifiable into three types in terms of molecular architecture, *i.e.*, proto type, chimera type, and tandem-repeat type.

The existence of these three types should have some important biological meaning because their binding functions are different. Therefore, it is important to study the properties of these two domains of the newly found tandem-repeat-type galectin of *C. elegans*. However, as the amount of protein available from *C. elegans* is too small to allow a variety of studies, we decided to apply gene technology for the production of larger amounts of the 32-kDa galectin. This approach also enables the production of two half molecules, each comprising one of the repeated galectin domains. Therefore, we established suitable conditions for expression of the whole molecule, and the N-terminal half (Nh) and C-terminal half (Ch) molecules in *Escherichia coli*, and purification procedures for each product. We found that all recombinant proteins had sugar-binding ability, but that there was a significant difference in the affinity for asialofetuin between Nh and Ch.

This suggests that the sugar-binding properties (binding strength and/or specificity) of the two lectin domains are not the same and that tandem-repeat-type galectins play a role as a heterobifunctional crosslinker.

MATERIALS AND METHODS

Expression and Purification of Recombinant Nematode Proteins—DNA fragments encoding 32 kDa-galectin (N32), its N-terminal half domain (Nh), and its C-terminal half domain (Ch) were amplified by PCR using cloned cDNA (*31*) as a template. Forward primers containing an *Nde*I site covering the initiation codon used for amplification were (5′-AACATATGTCCGCCGAAGAGCCA-3′) for N32 and Nh, and (5′-TTCACTGGGGACATATGTACTACCCA-3′) for Ch. Reverse primers were (5′-TTGGATCCTCAGTGAACGTGGTTCAA-3′) containing a *Bam*HI site downstream of the desired stop codon for Nh and a pBluescript-specific oligonucleotide (5′-GTTTTCCCAGTCACGAC-3′) for N32 and Ch. PCR products were ligated to pCR II, which was then employed for transformation of competent *E. coli* INVαF′ by the use of a TA cloning kit (Invitrogen). The inserts were excised from the clones with *Nde*I and *Xho*I (N32 and Ch) or with *Nde*I and *Bam*HI (Nh). Excised fragments were ligated into *Nde*I/*Xho*I (N32 and Ch)- or *Nde*I/*Bam*HI (Nh)-digested pET 21a (Novagen) and used to transform *E. coli* HB101 cells. To avoid obtaining the pCR II transformant, we digested the ampicillin-resistant region of pCR II with *Ava*II. Obtained pET plasmids having inserts coding for N32, Nh, or Ch were incorporated into *E. coli* BL21(DE3) cells. Positive clones screened by plasmid minipreps and restriction enzyme analysis were grown at 37°C in 2×YT medium containing 50 μg/ml ampicillin until the absorbance at 600 nm reached 0.6. Isopropyl-β-thiogalactopyranoside (IPTG) was added to give a final concentration of 1 mM. After an additional 2-h culture period, the cells were harvested and disrupted in EDTA-ME-PBS (20 mM Na-phosphate, pH 7.2, containing 150 mM NaCl, 4 mM β-mercaptoethanol, and 2 mM EDTA) by sonication as described previously (*35*). The debris was centrifuged at 14,000×*g* for 30 min. Concentrated recombinant proteins forming inclusion bodies were solubilized in 8 M urea in EDTA-ME-PBS, then allowed to refold by gradual removal of the urea by dialysis. The refolded recombinant protein solution was applied to an asialofetuin-Sepharose 4B column [column volume, 10 ml; 9.1 mg protein/ml gel prepared according to the method of de Waard *et al.* (*36*)] that had previously been equilibrated with EDTA-ME-PBS. After the column had been washed, the adsorbed proteins were eluted with 0.1 M lactose in EDTA-ME-PBS. Affinity-purified recombinant protein solution was dialyzed extensively against 1/5 PBS(−) to remove lactose, and the protein was freeze-dried and stored at −80°C until used. Since the affinity of Nh for asialofetuin-Sepharose 4B was very weak, the refolded Nh solution was applied to DEAE-Toyopearl 650S (Tosoh, Tokyo) equilibrated with 10 mM Tris-HCl pH 8.0. An NaCl gradient was applied from 0 mM NaCl to 100 mM NaCl in 10 mM Tris-HCl pH 8.0. Since Nh was already concentrated by the formation of inclusion bodies, this anion-exchange chromatography step was very effective for obtaining pure Nh. After extensive dialysis against 1/5 PBS(−), purified Nh was freeze-dried and then stored at −80°C until used.

DNA Sequencing of the Plasmid Construct for Expression—DNA sequences of inserts coding for N32, Nh, and Ch in pET plasmids were determined by the dye-termination cycle sequencing method with an Applied Biosystems DNA sequencer 373S according to the manufacturer's instructions. Primers used for sequencing were two pET 21a vector-specific primers, *i.e.*, T3 promoter primer and T7 promoter primer, and several inner primers designed according to the cDNA sequence of the 32-kDa galectin (*31*).

Hemagglutination Assay—Hemagglutination activity was measured according to Barondes *et al.* (*37*). A sample of 25 μl of recombinant protein was serially twofold diluted in a 96-well microtiter V plate (Greiner), and 25 μl each of PBS, 1% (w/v) bovine serum albumin in 0.15 M NaCl, and 4% (v/v) glutaraldehyde-fixed trypsin-treated rabbit erythrocytes in PBS(−) were added. After standing for 60 min at room temperature, the end-point showing the minimum concentration giving hemagglutination was determined.

High-Performance Gel Permeation Chromatography—Molecular weights of the recombinant proteins under non-denaturing conditions were estimated by high-performance gel permeation chromatography on a TSKgel G2000SWXL column (7.8×300 mm; Tosoh) at a flow rate of 1 ml/min in EDTA-PBS [20 mM Na-phosphate (pH 7.2), 0.15 M NaCl, 2 mM EDTA]. The amount of recombinant proteins applied to the column was 1 μg. Elution of protein was monitored by measuring the fluorescence at 350 nm (ex-

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

citation at 280 nm). Bovine serum albumin (66 kDa), ovalbumin (45 kDa), soybean trypsin inhibitor (20 kDa), and β-lactoglobulin (18 kDa) were used as molecular weight markers.

Frontal Affinity Chromatography on Asialofetuin-Sepharose—Frontal affinity chromatography was performed as previously described (*38*). A glass column (column volume, 2 ml) was packed with asialofetuin-Sepharose 4B. All the chromatographic operations were carried out at 4°C. After one operation, the column was washed with 10 ml of 0.1 M lactose in EDTA-ME-PBS in order to remove lectin proteins completely. Then, for re-equilibration, more than 300 ml of EDTA-ME-PBS was applied. Recombinant proteins were dissolved in EDTA-ME-PBS to make 5 μg/ml and applied to the column continuously at the flow rate of 0.3 ml/min; and fractions of 5 ml (N32, Ch) or 1 ml (Nh or recombinant proteins in the presence of lactose) were collected. To determine the precise elution volume of the front of the applied protein (V) from the elution profile (curve I in Fig. 1), we employed the following equation:

$$V = na - a\frac{\sum_{i=1}^{n}[\mathrm{A}]_i}{[\mathrm{A}]_0} \quad (1)$$

where a is the volume of one fraction, n is the tube number of a certain fraction at the plateau, and $[\mathrm{A}]_i$ is the concentration of fraction i. Although the V value includes the volume of the tubing from the outlet of the column to the fraction collector, this can be neglected because we always consider values relative to V_0, which is the elution volume in the case without specific interaction (curve II in Fig. 1).

Protein Concentration Determination and SDS-PAGE—Protein concentration in each fraction was determined by the method of Bradford (*39*). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions was performed by the method of Laemmli (*40*) using 14% gels. Protein bands were visualized by silver-staining with a Silver Stain Kit Wako (Wako Pure Chemicals, Tokyo).

Western Blotting—Protein solutions were mixed with Laemmli's sample buffer (*40*) and boiled for 3 min, subjected to SDS-PAGE on a 14% polyacrylamide slab gel and then transferred electrophoretically to a nitrocellulose filter. The filter was incubated with anti-galectin antiserum (1:1,000 dilution) in Tris-buffered saline (50 mM Tris-HCl, 150 mM NaCl, pH 7.5) containing 0.1% (w/v) Tween 20 (TBS-T). Antiserum raised against purified *C. elegans* galectin was prepared as described before (*30*). Western blotting experiments showed that this antiserum recognizes only the 32-kDa galectin in crude extracts of *C. elegans*. As molecular weight markers, bovine serum albumin, ovalbumin, carbonic anhydrase, β-lactoglobulin, and lysozyme were used. The part of the nitrocellulose filter bearing the transferred marker proteins was cut out and stained with 0.25% (w/v) amido black in 45% (v/v) acetic acid + 10% (v/v) methanol. The remaining part of the filter was incubated with anti-galectin antiserum, washed three times with TBS-T, and incubated with horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (1: 1,000 dilution, purchased from Seikagaku, Tokyo). The filter was washed again with TBS-T three times, and the signals were detected by a chromogenic detection system for HRP (POD immunostain set, Wako Pure Chemicals, Osaka).

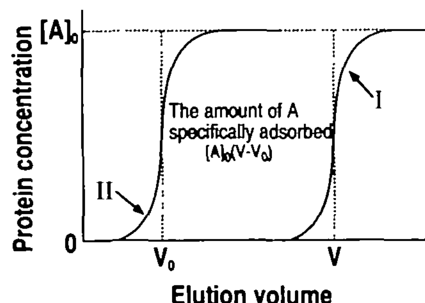

Fig. 1. **Elution profiles in frontal affinity chromatography.** Curve I: Elution profile of an analyte, in the present case one of the recombinant lectins that specifically interacts with the immobilized ligand (elution volume = V). Curve II: Elution profile of the analyte when the specific interaction with the ligand is abolished in the presence of excess free ligands, in this case lactose (elution volume = V_0).

Sequencing the N-Terminal Part of the Recombinant Protein—Sequencing of the N-terminal portion of the recombinant proteins by the use of an Applied Biosystems pulsed-liquid phase sequencer (model 477A) was performed essentially as described previously (*6, 9*).

Determination of the Molecular Mass of the Recombinant Proteins by Matrix-Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) Mass Spectrometry—Recombinant proteins were dialyzed against distilled water and freeze-dried. They were dissolved in 0.1% trifluoroacetic acid to make 10 pmol/μl. α-Cyano-4-hydroxycinnamic acid (CHCA) was used as the matrix. One microliter of the protein solution and one microliter of the saturated matrix solution were applied to a stainless steel probe tip and dried at room temperature. The molecular masses of the recombinant proteins were determined using a Shimadzu Compact MALDI IV mass spectrometer.

Circular Dichroism (CD) Spectra of Recombinant N32, Nh, and Ch—CD spectra of recombinant proteins at the concentration of 0.25 mg/ml in PBS(−) were obtained with a Jasco J-600 spectropolarimeter.

RESULTS

Production of Recombinant Proteins—Recombinant 32-kDa galectin of *C. elegans* (N32), its N-terminal half lectin domain (Nh), and its C-terminal half lectin domain (Ch) were expressed in *E. coli* as described under "MATERIALS AND METHODS." When the IPTG-induced bacteria were lysed, most of the recombinant proteins precipitated as inclusion bodies. After solubilization of the inclusion bodies with 8 M urea, recombinant proteins were allowed to refold by the slow removal of the urea by dialysis. Since a large amount of the recombinant protein was produced in *E. coli* cells and precipitated from the cytosol, N32, Nh, or Ch was already the major component of the solution. N32 and Ch were then purified by affinity chromatography on asialofetuin-Sepharose 4B. Figure 2A shows a typical elution profile of N32. Most of the *E. coli* proteins solubilized by urea were not adsorbed on the column. After the column had been washed, specifically adsorbed N32 was eluted with 0.1 M lactose. As shown in Fig. 2B, in the lactose-eluted fractions (fraction numbers 19 to 24), the recombinant N32 protein was detected as the major component with a minor

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

fraction of proteins that seemed to be the degradation products. A portion of N32 was eluted from the column before the addition of lactose (fraction numbers 12 to 16). This is probably because the affinity of N32 for the immobilized asialofetuin is not strong; thus, the adsorbed N32 started to be released during the washing step. Recombinant Ch was also purified by the same procedure.

Since Nh had a very weak affinity for asialofetuin-Sepharose 4B, retarded fractions eluted right after the flow-through fraction were collected in the early stages of our experiments. However, the yield was low. Therefore, we examined ion-exchange chromatography as described in "MATERIALS AND METHODS." The solubilized and renatured Nh solution was applied to DEAE-Toyopearl 650 S, and the protein was eluted with a gradient of NaCl concentration (10 mM Tris-HCl, pH 7.2, to 10 mM Tris-HCl+100 mM NaCl, pH 7.2). This procedure proved to be very effective, because only the recombinant Nh was eluted and the other proteins remained adsorbed on the column.

For each recombinant protein the yield was 10 to 20 mg from a 1-liter culture. The N-terminal amino acid sequences of all three recombinant proteins determined by a gas phase protein sequencer were consistent with those expected. In fact, N-terminal Met was removed from N32 and Nh, probably by methionine aminopeptidase in *E. coli*, which cleaves the initial Met residue preceding the Ser residue (*41*). The DNA sequences of inserts coding N32, Nh, and Ch in pET plasmids were identical to those of the corresponding regions of the cDNA of the 32-kDa galectin (*31*). The molecular weights of the three recombinant proteins were determined by MALDI-TOF mass spectrometry. The observed $[M+H]^+$ ion peaks of N32, Nh, and Ch were at m/z 31,493, 16,225, and 15,213, respectively. Peaks corresponding to degradation products were not detected (data not shown). The observed molecular mass values were consistent with those expected: N32 (31,678), Nh (16,204), and Ch (15,193). Though a relatively large difference was observed in the case of N32, it seemed to be within experimental error because the size of the protein was close to the limit of the equipment. Therefore, each recombinant protein seemed to have been completely expressed and no appreciable degradation occurred.

The purity of each recombinant protein was checked by SDS-PAGE (Fig. 3A). Each of recombinant N32, Nh, and Ch showed one major band approximately at the position expected from the deduced molecular weight. Though Nh migrated faster than Ch, this does not mean that the expressed Nh was truncated, because the molecular mass was confirmed by MALDI-TOF mass spectrometry. The recombinant proteins were also detected by Western blotting using the anti-32-kDa galectin antiserum raised against the native *C. elegans* 32-kDa galectin (Fig. 3B).

Analysis of the Secondary Structures of the Recombinant Proteins by CD Spectroscopy—The CD spectra of the recombinant proteins dissolved in PBS(−) are shown in Fig. 4. They all showed a typical β-sheet structure profile (Fig. 4, a-c) resembling that of human galectin-1 (C2S; Fig. 4d) and thus the proteins seem to have refolded properly after the dialysis to remove the urea that had been used to solubilize the inclusion bodies.

Estimation of Molecular Weight by Gel Filtration—Proto-type mammalian galectins are known to form homodimers, to show hemagglutination activity, and to be eluted at the position of a dimer on gel filtration chromatography. To test if Nh or Ch, which are homologous to mammalian proto-type galectins, forms a dimer and if N32 forms oligomers, we applied these recombinant proteins to a high-performance gel filtration column (Tosoh TSKgel G2000-SWXL). The eluted proteins were monitored by measuring the fluorescence at 350 nm (excitation at 280 nm). As shown in Fig. 5, N32, Nh, and Ch each gave only a single peak at the position of the corresponding monomer. The estimated molecular weight of Ch was rather smaller than expected, possibly because it had a weak interaction with the resin.

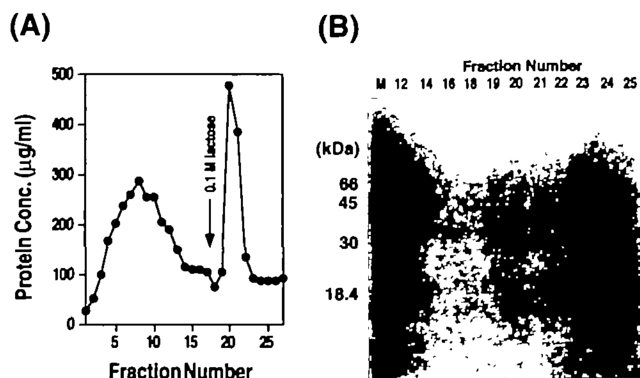

Fig. 2. **Purification of recombinant N32 by affinity chromatography on asialofetuin-Sepharose 4B.** (A) Elution profile of recombinant N32. Recombinant N32 protein which precipitated as inclusion bodies was solubilized and renatured as described in "MATERIALS AND METHODS." N32 was then applied to an asialofetuin-Sepharose 4B column. After the column had been washed with EDTA-ME-PBS, specifically adsorbed N32 was eluted with EDTA-ME-PBS containing 0.1 M lactose. (B) Analysis by SDS-PAGE under reducing conditions. Eluted fractions from the asialofetuin-Sepharose 4B were subjected to SDS-PAGE, and protein was visualized by silver staining. Numbers indicated on the gel correspond to the fraction number of asialofetuin-Sepharose column eluate in "A."

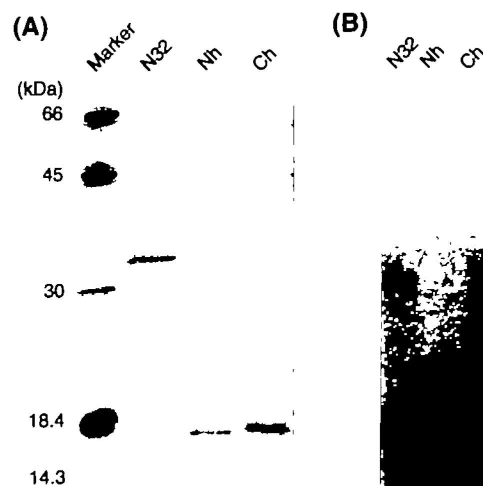

Fig. 3. **Purities of recombinant N32, Nh, and Ch.** Purified recombinant N32, Nh, and Ch were subjected to SDS-PAGE and visualized by silver staining (A). They were also subjected to immunoblotting analysis (B), and all three proteins showed immunoreactivity against antiserum raised against native 32-kDa galectin purified from *C. elegans*.

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

The mixture of Nh and Ch did not form heterodimers; instead, two peaks corresponding to each monomer protein were observed (Fig. 5A, d). Therefore, all three recombinant proteins proved to exist as monomers.

Hemagglutination Activity—All three recombinant proteins were subjected to a hemagglutination assay. N32 had a very strong hemagglutination activity, and the minimum concentration required for hemagglutination was 0.5 μg/ml. On the other hand, Nh or Ch had almost no hemagglutination activity, even at the concentration of 100 μg/ml. Because N32 is a combination of Nh and Ch, the mixture of the latter two was also tested, but no hemagglutination activity was observed. These results are consistent with those of gel filtration. Only N32, which has two lectin domains, could crosslink erythrocytes. However, neither Nh nor Ch can form a dimer and they remain monovalent. Therefore, they can not crosslink erythrocytes.

Estimation of Binding Strength of the Recombinant Proteins by Frontal Affinity Chromatography—Relative sugar binding strength was compared by frontal affinity chromatography. A solution of 5 μg/ml ($=[A]_0$) recombinant protein in EDTA-ME-PBS was applied continuously to asialofetuin-Sepharose 4B (bed volume, 2 ml) equilibrated with the same buffer. The elution volume of the front of each recombinant protein was measured as described in "MATERIALS AND METHODS." The relation between $[A]_0$ and V is given by the following equation (*38*),

$$V-V_0=\frac{v[B]_0}{K_d+[A]_0} \quad (2)$$

where V_0 is the elution volume of the front in the case without specific interaction (*e.g.*, in the presence of 0.1 M lactose), $v[B]_0$ is the total amount of immobilized ligand (saccharide chains of asialofetuin that can interact with the applied protein), and K_d is the dissociation constant. The smaller the $[A]_0$ value, the larger the V value. If $[A]_0 \ll K_d$, *i.e.*, if $[A]_0$ is negligibly small compared with K_d, $V-V_0$ approaches the maximum value, V_m, which is independent of $[A]_0$, and the following equation is obtained:

$$V_m=\frac{v[B]_0}{K_d} \quad (3)$$

This means that the value of V_m is proportional to the affinity of the recombinant protein for asialofetuin-Sepharose. Therefore, if we measure elution volumes for the recombinant proteins at an adequately low concentration, these values are inversely proportional to the dissociation constant (or proportional to the association constant).

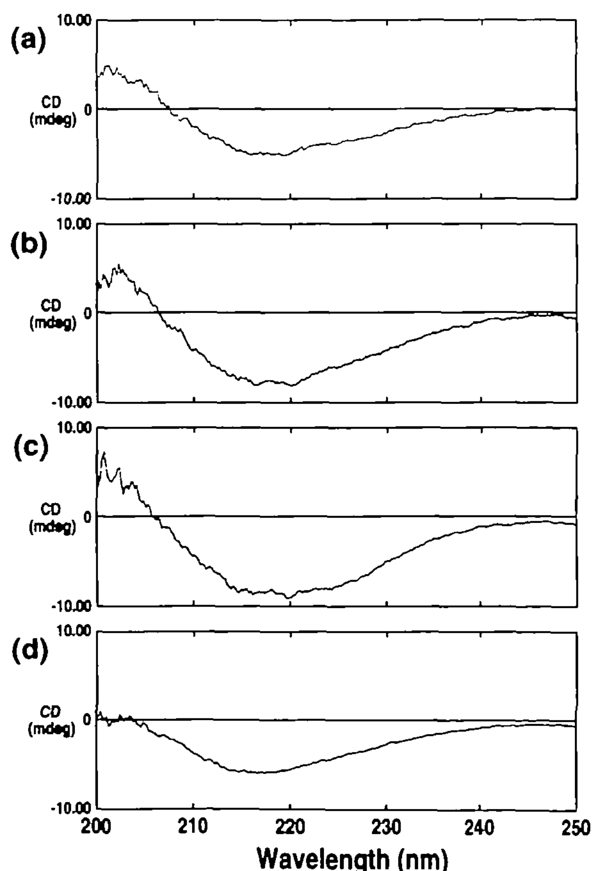

Fig. 4. **CD spectra of recombinant N32, Nh, and Ch.** CD spectra of recombinant N32 (a), Nh (b), Ch (c), and recombinant human galectin C2S (d) were obtained at the concentration of 0.25 mg/ml. All four spectra indicate an abundance of β-structure.

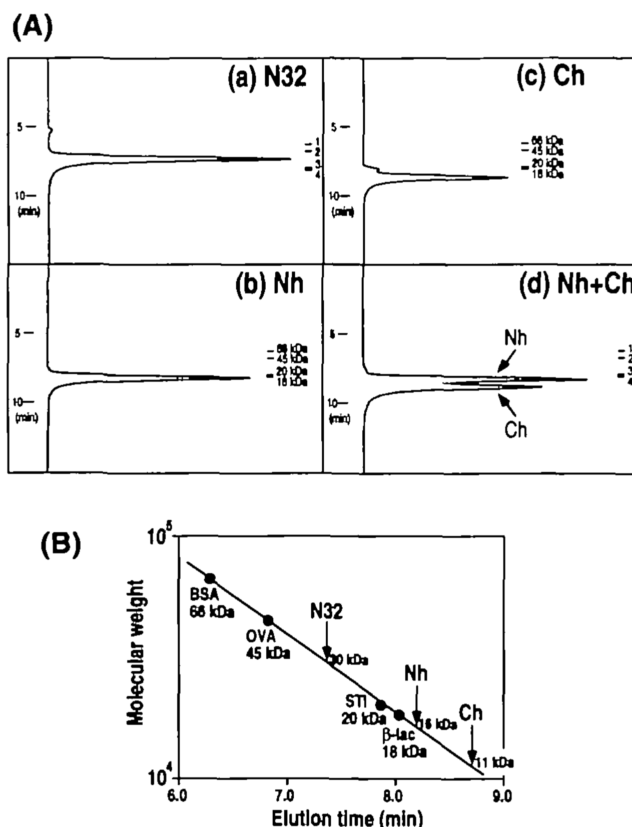

Fig. 5. **Gel permeation chromatography of recombinant proteins.** (A) Elution profile of recombinant proteins. Recombinant N32 (a), Nh (b), Ch (c), and a mixture of Nh and Ch (d) were applied to a TSKgelG2000SWXL column. The molecular weight markers used were: 1, bovine serum albumin (66 kDa); 2, ovalbumin (45 kDa); 3, soybean trypsin inhibitor (20 kDa); and 4, β-lactoglobulin (18 kDa). (B) Retention times of recombinant N32, Nh, and Ch compared with those of the molecular weight markers. Molecular weights of the recombinant proteins were estimated from their retention times.

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

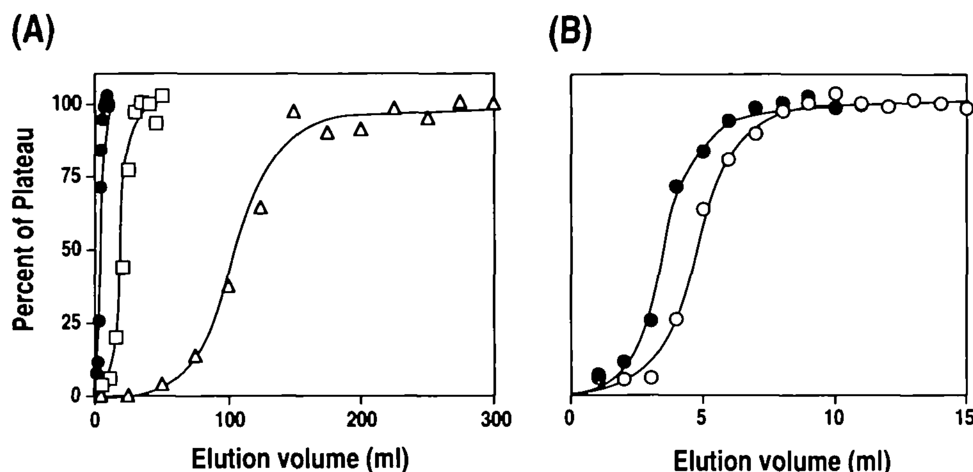

Fig. 6. **Frontal affinity chromatography of recombinant proteins on asialofetuin-Sepharose 4B.** (A) Elution profiles of N32 and Ch. Solutions of N32 (△) and Ch (□) (5 μg/ml each) were continuously applied to a 2-ml asialofetuin-Sepharose 4B column. Eluted protein concentration of each fraction was measured. Closed circles (●) indicate the elution profile of N32 in the presence of 0.1 M lactose. (B) Elution profile of Nh. A solution of Nh (5 μg/ml) was applied to the same column in the absence (○) or presence (●) of 0.1 M lactose.

The elution profiles of N32, Nh, and Ch are shown in Fig. 6A. The calculated values of V_m of each recombinant protein were 100, 1.14, and 14.6 ml for N32, Nh, and Ch, respectively. Ch had much weaker affinity for asialofetuin than N32 (7 times lower). Nh proved to have very low affinity (100 times lower than N32), but did retain some affinity since the elution front of Nh was slightly retarded (Fig. 6B).

Though the primary structures of both of the repeated domains of *C. elegans* 32-kDa galectin are homologous to those of the vertebrate lectins, some important residues, which had been elucidated by site-directed mutagenesis studies (*3, 42, 43*) and X-ray crystallography (*44-46*), were found to be replaced. In the Nh region the asparagine residue (Asn-61), conserved in almost all galectins, had been substituted by serine (Ser-61). Conserved amino acids on both sides of this asparagine had also been changed to valine (conserved sequence His-Phe-Asn-Pro-Arg-Phe was changed to His-Val-Ser-Val-Arg-Phe). This might greatly affect the binding properties of the Nh domain, because this highly conserved asparagine was reported to form a hydrogen bond with the C-4 hydroxyl group of the galactose moiety in lactose (*44*). This interaction seems to be crucial for the recognition of galactose. This substitution might explain why Nh showed very weak affinity for asialofetuin-Sepharose.

To determine whether this amino acid substitution is the cause of the weakness of the affinity of Nh for asialofetuin, we conducted site-directed mutagenesis studies by PCR. The amino acids His-Val-Ser-Val-Arg-Phe of Nh were replaced by His-Phe-Asn-Pro-Arg-Phe. However, no significant change in the binding strength of the mutagenized Nh was observed (data not shown).

DISCUSSION

The vector pET 21a used for the expression of the recombinant proteins was very effective. Though the produced proteins precipitated as inclusion bodies, they could be solubilized with 8 M urea and renatured by dialysis. Purified N32 and Ch were obtained by affinity chromatography. Since the affinity of Nh to asialofetuin-Sepharose was very weak, Nh was purified by ion-exchange chromatography (DEAE-Toyopearl 650S). Based on the CD spectra of the recombinant proteins, they all seem to have refolded properly and be rich in β-structure [as reported for human galectin-2 (*44*) and bovine galectin-1 (*45*)].

The recombinant galectin domains separately produced by *E. coli* did not form homodimers (Fig. 5), in contrast to mammalian proto-type galectins. They did not form heterodimers either.

Since neither recombinant Nh nor Ch formed homodimers, neither could crosslink erythrocytes and each of them was thus inactive in the hemagglutination assay. Even the mixture of Nh and Ch did not form dimers, and as a result, showed no hemagglutination activity. It is noteworthy that the recombinant N32 had a strong hemagglutination activity towards trypsin-treated glutaraldehyde-fixed rabbit erythrocytes, though N32 itself seems to exist as a monomer. The hemagglutination titer was almost the same as that of human galectin-1 (proto-type) or stronger. This strongly suggests that both of the carbohydrate-recognition domains of N32 have binding activities toward the saccharide chains on the surface of erythrocytes, resulting in the crosslinking of erythrocytes. Measurement of the hemagglutination activity of the native 32-kDa galectin was not possible because of the poor yield from *C. elegans*. Therefore, it is possible that the characteristics of the recombinant galectin did not exactly reflect those of the native galectin. However, such an approach was successful in the case of the human galectin because the hemagglutination activity of the recombinant human 14 kDa galectin was almost the same as that of the native galectin purified from placenta (*35*). It was reported that recombinant rat galectin-8 showed hemagglutination activity (*25*), though it is not certain if the recombinant protein forms an oligomeric structure.

The results of frontal affinity chromatography showed that the affinities of recombinant N32, Nh, and Ch were completely different. It is reasonable that N32 had the strongest affinity to asialofetuin-Sepharose 4B, because it has two binding sites and thus a higher probability of interacting with the ligands on the column. If the two binding sites have the same binding strength, the apparent binding strength of the whole molecule should be twice that of a single binding site. However, the results were more complicated. The affinities of Ch and Nh were one-seventh and one-hundredth of that of N32, respectively. The

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

binding strength of N32 cannot be explained by the combination of those of Nh and Ch. It is possible that the binding sites of both half molecules cannot assume a fully active state if they are separated, or that refolding of them was not perfect. The possibility that Nh is not refolded properly after urea solubilization and dialysis can be excluded, since the CD spectrum of Nh (Fig. 4b) showed a typical β-sheet structure profile like that of human galectin-1 (C2S; Fig. 4d). Recombinant Nh seems to have refolded properly, as well as N32 or C2S. Furthermore, the molecular masses of the recombinant Nh and Ch determined by MALDI-TOF mass spectrometry strongly suggested that their entire polypeptide chains were expressed. It is not yet known whether the spatial arrangement of the two carbohydrate binding sites of N32 resembles that of the homodimer of the proto-type galectins such as galectin-1 (*45*) and galectin-2 (*44*). Further studies, especially from the standpoint of three-dimensional structure, are needed to obtain definite answers, and such studies are in preparation. It is possible that Ch and Nh are different in not only binding strength, but also binding specificity, though we have not yet found saccharides having strong affinity for Nh. When N32 crosslinks erythrocytes, these two sites probably bind to different structures on the surface of the erythrocytes. In that case, tandem-repeat type galectin would act as a heterobifunctional crosslinker protein.

Mammalian galectins containing two tandemly repeated lectin domains are galectin-4 from rat (*20*) and pig (*21*) and galectin-8 from rat (*25*). Their two lectin domains are separated by a linking peptide composed of 25 to 32 amino acids, whereas *C. elegans* 32-kDa galectin has no linking peptide. Sugar-binding properties of the two lectin domains of rat galectin-4 have been reported. The recombinant N-terminal half-domain (domain I) and native C-terminal half-domain (domain II; a proteolytic product of native galectin-4, which is the main component after the purification procedures on lactosyl-Sepharose chromatography) were found to have similar binding specificities.

In rat galectin-4, most of the important amino acids in both of the carbohydrate recognition domains are conserved. However, in domain-I (corresponding to the Nh domain of *C. elegans* 32-kDa galectin), the arginine residue that forms hydrogen bonds with the C-2 and C-3 hydroxyl groups of the glucose moiety in lactose is substituted to Lys-89 (*20*). This might be the cause of the slight difference observed in the binding strength of the two lectin domains. However, this substitution is between basic amino acids, so the change in the affinity should be much smaller than that observed in the case of the nematode 32-kDa galectin. The binding properties of the two lectin domains in porcine galectin-4 and rat galectin-8 have not been studied so far.

Since the tandemly repeated 32-kDa galectin of *C. elegans* has a strong hemagglutination activity, though the binding strengths of its two lectin domains with asialofetuin-Sepharose are different, we suggest that the 32-kDa galectin has a role to crosslink structurally different sugar chains of glycoconjugates of the nematode. Identification of different ligands recognized by the two lectin domains is important. Candidate endogenous glycoconjugate ligands for the recently identified proto-type 16-kDa galectin of *C. elegans* (*47*) have been isolated from the Triton X-100 extract of the worm by means of affinity chromatography using immobilized recombinant 16-kDa galectin (*47*). Similar glycoproteins have also been detected by the same procedure using N32-Sepharose. Further characterization of these glycoproteins, including the structural analysis of the carbohydrate chains, might solve the problem. The carbohydrate structure that is recognized by Nh might be considerably different from the lactosamine structure since the affinity of Nh for asialofetuin-Sepharose is relatively weak in comparison with that of N32 or Ch.

This 32-kDa galectin was found to be localized most abundantly in the adult cuticle of *C. elegans* by immunohistochemical studies (*48*). Therefore, it may be an essential component of the adult cuticular matrix as a heterobifunctional crosslinker for the construction of the tough and durable outer barrier of the worm body. The lectin was also found in the terminal bulb of the pharynx, where the food is concentrated and processed before being pumped into the gut. Although its precise role remains to be determined, the heterobifunctional galectin seems to have important but different roles in different parts of the worm.

We are very grateful to Professor Ichio Shimada, Dr. Chiaki Nishimura, Dr. Koichi Kato, and Mr. Yoshiki Yamaguchi of the Faculty of Pharmaceutical Sciences, The University of Tokyo, for the use of the circular dichroism and MALDI-TOF mass spectrometry facilities. We are grateful to the *Caenorhabditis* Genetics Center, which is supported by the National Institutes of Health National Center for Research Resources, for kindly providing wild-type *C. elegans*. We also thank M. Yano, N. Umemoto, M. Ishizuna, T. Narita, H. Zen-in, I. Hayashi, and A. Nakano for technical assistance.

REFERENCES

1. Drickamer, K. (1988) Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem.* **263**, 9557-9560
2. Hirabayashi, J. (1993) A general comparison of two major families of animal lectins. *Trends Glycosci. Glycotechnol.* **5**, 251-270
3. Hirabayashi, J. and Kasai, K. (1993) The family of metazoan metal-independent β-galactoside-binding lectins: structure, function and molecular evolution. *Glycobiology* **3**, 297-304
4. Barondes, S.H., Cooper, D.N.W., Gitt, M.A., and Leffler, H. (1994) Galectins. Structure and function of a large family of animal lectins. *J. Biol. Chem.* **269**, 20807-20810
5. Kasai, K. and Hirabayashi, J. (1996) Galectins: A family of animal lectins that decipher glycocodes. *J. Biochem.* **119**, 1-8
6. Hirabayashi, J., Kawasaki, H., Suzuki, K., and Kasai, K. (1987) Complete amino acid sequence of 14 kDa β-galactoside-binding lectin of chick embryo. *J. Biochem.* **101**, 775-787
7. Ohyama, Y., Hirabayashi, J., Oda, Y., Ohno, S., Kawasaki, H., Suzuki, K., and Kasai, K. (1986) Nucleotide sequence of chick 14K β-galactoside-binding lectin mRNA. *Biochem. Biophys. Res. Commun.* **134**, 51-56
8. Barondes, S.H., Castronovo, V., Cooper, D.N.W., Cummings, R.D., Drickamer, K., Feizi, T., Gitt, M.A., Hirabayashi, J., Hughes, C., Kasai, K., Leffler, H., Liu, F.-T., Lotan, R., Mercurio, A.M., Monsigny, M., Pillai, S., Poirer, F., Raz, A., Rigby, P.W., Rini, J.M., and Wang, J.L. (1994) Galectins: A family of animal β-galactoside-binding lectins. *Cell* **76**, 597-598
9. Hirabayashi, J. and Kasai, K. (1988) Complete amino acid sequence of a β-galactoside-binding lectin from human placenta. *J. Biochem.* **104**, 1-4
10. Hirabayashi, J., Ayaki, H., Soma, G., and Kasai, K. (1989) Cloning and nucleotide sequence of a full-length cDNA for human 14 kDa β-galactoside-binding lectin. *Biochim. Biophys. Acta* **1008**, 85-91
11. Clerch, L.B., Whitney, P., Hass, M., Brew, K., Miller, T., Werner, R., and Massaro, D. (1988) Sequence of a full-length cDNA for rat lung β-galactoside-binding protein: Primary and

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

secondary structure of the lectin. *Biochemistry* **27**, 692-699
12. Wells, V. and Mallucci, L. (1991) Identification of an autocrine negative growth factor: Mouse β-galactoside-binding protein is a cytostatic factor and cell growth regulator. *Cell* **64**, 91-97
13. Abbott, W.M., Mellor, A., Edwards, Y., and Feizi, T. (1989) Soluble bovine galactose-binding lectin. cDNA cloning reveals the complete amino acid sequence and an antigenic relationship with the major encephalitogenic domain of myelin basic protein. *Biochem. J.* **259**, 283-290
14. Gitt, M.A., Massa, S.M., Leffler, H., and Barondes, S.H. (1992) Isolation and expression of a gene encoding L-14-II, a new human soluble lactose-binding lectin. *J. Biol. Chem.* **267**, 10601-10606
15. Robertson, M.W., Albrandt, K., Keller, D., and Liu, F.-T. (1990) Human IgE-binding protein: A soluble lectin exhibiting a highly conserved interspecies sequence and differential recognition of IgE glycoforms. *Biochemistry* **29**, 8093-8100
16. Cherayil, B.J., Chaitovitz, S., Wong, C., and Pillai, S. (1990) Molecular cloning of a human macrophage lectin specific for galactose. *Proc. Natl. Acad. Sci. USA* **87**, 7324-7328
17. Oda, Y., Leffler, H., Sakakura, Y., Kasai, K., and Barondes, S.H. (1991) Human breast carcinoma cDNA encoding a galactose-binding lectin homologous to mouse Mac-2 antigen. *Gene* **99**, 279-283
18. Albrandt, K., Orida, N.K., and Liu, F.-T. (1987) An IgE-binding protein with a distinctive repetitive sequence and homology with an IgG receptor. *Proc. Natl. Acad. Sci. USA* **84**, 6859-6963
19. Jia, S. and Wang, J.L. (1988) Carbohydrate binding protein 35. Complementary DNA sequence reveals homology with proteins of the heterogeneous nuclear RNP. *J. Biol. Chem.* **263**, 6009-6011
20. Oda, Y., Herrmann, J., Gitt, M.A., Turck, C.W., Burlingame, A.L., Barondes, S.H., and Leffler, H. (1993) Soluble lactose-binding lectin from rat intestine with two different carbohydrate-binding domains in the same peptide chain. *J. Biol. Chem.* **268**, 5929-5939
21. Chiu, M.L., Parry, D.A.D., Feldman, S.R., Klapper, D.G., and O'Keefe, E.J. (1994) An adherens junction protein is a member of the family of lactose-binding lectins. *J. Biol. Chem.* **269**, 31770-31776
22. Gitt, M.A., Wiser, M.F., Leffler, H., Herrmann, J., Xia, Y.-R., Massa, S.M., Cooper, D.N.W., Lusis, A.J., and Barondes, S.H. (1995) Sequence and mapping of galectin-5, a β-galactoside-binding lectin, found in rat erythrocytes. *J. Biol. Chem.* **270**, 5032-5038
23. Madsen, P., Rasmussen, H.H., Flint, T., Gromov, P., Kruse, T.A., Honoré, B., Vorum, H., and Celis, J.E. (1995) Cloning, expression, and chromosome mapping of human galectin-7. *J. Biol. Chem.* **270**, 5823-5829
24. Magnaldo, T., Bernerd, F., and Darmon, M. (1995) Galectin-7, a human 14-kDa S-lectin, specifically expressed in keratinocytes and sensitive to retinoic acid. *Dev. Biol.* **168**, 259-271
25. Hadari, Y.R., Paz, K., Dekel, R., Mestrovic, T., Accili, D., and Zick, Y. (1995) Galectin-8. A new rat lectin, related to galectin-4. *J. Biol. Chem.* **270**, 3447-3453
26. Sakakura, Y., Hirabayashi, J., Oda, Y., Ohyama, Y., and Kasai, K. (1990) Structure of chicken 16-kDa β-galactoside-binding lectin. Complete amino acid sequence, cloning of cDNA, and production of recombinant lectin. *J. Biol. Chem.* **265**, 21573-21579
27. Paroutaud, P., Levi, G., Teichberg, V.I., and Strosberg, A.D. (1987) Extensive amino acid sequence homologies between animal lectins. *Proc. Natl. Acad. Sci. USA* **84**, 6345-6348
28. Marschal, P., Herrmann, J., Leffler, H., Barondes, S.H., and Cooper, D.N.W. (1992) Sequence and specificity of a soluble lactose-binding lectin from *Xenopus laevis* skin. *J. Biol. Chem.* **267**, 12942-12949
29. Ozeki, Y., Matsui, T., Nitta, K., Kawauchi, H., Takayanagi, Y., and Titani, K. (1991) Purification and characterization of β-galactoside binding lectin from frog (*Rana catesbeiana*) eggs. *Biochem. Biophys. Res. Commun.* **178**, 407-413
30. Hirabayashi, J., Satoh, M., Ohyama, Y., and Kasai, K. (1992) Purification and characterization of β-galactoside-binding proteins from *Caenorhabditis elegans*. *J. Biochem.* **111**, 553-555
31. Hirabayashi, J., Satoh, M., and Kasai, K. (1992) Evidence that *Caenorhabditis elegans* 32-kDa β-galactoside-binding protein is homologous to vertebrate β-galactoside-binding lectins. cDNA cloning and deduced amino acid sequence. *J. Biol. Chem.* **267**, 15485-15490
32. Pfeifer, K., Haasemann, M., Gamulin, V., Bretting, H., Fahrenholz, F., and Müller, W.E.G. (1993) S-type lectins occur also in invertebrates: High conservation of the carbohydrate recognition domain in the lectin genes from the marine sponge *Geodia cydonium*. *Glycobiology* **3**, 179-184
33. Brenner, S. (1974) The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71-94
34. Wood, W.B. (1988) *The Nematode Caenorhabditis elegans*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
35. Hirabayashi, J., Ayaki, H., Soma, G., and Kasai, K. (1989) Production and purification of a recombinant human 14 kDa β-galactoside-binding lectin. *FEBS Lett.* **250**, 161-165
36. de Waard, A., Hickman, S., and Kornfeld, S. (1976) Isolation and properties of β-galactoside binding lectins of calf heart and lung. *J. Biol. Chem.* **251**, 7581-7587
37. Nowak, T.P., Haywood, P.L., and Barondes, S.H. (1976) Developmentally regulated lectin in embryonic chick muscle and a myogenic cell line. *Biochem. Biophys. Res. Commun.* **68**, 650-657
38. Kasai, K., Oda, Y., Nishikata, M., and Ishii, S. (1986) Frontal affinity chromatography: Theory for its application to studies on specific interactions of biomolecules. *J. Chromatogr.* **376**, 33-47
39. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254
40. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
41. Ben-Bassat, A., Bauer, K., Chang, S.-Y., Myambo, K., Boosman, A., and Chang, S. (1987) Processing of the initiation methionine from proteins: properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. *J. Bacteriol.* **169**, 751-757
42. Hirabayashi, J. and Kasai, K. (1991) Effect of amino acid substitution by site-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa β-galactoside-binding lectin. *J. Biol. Chem.* **266**, 23648-23653
43. Hirabayashi, J. and Kasai, K. (1994) Further evidence by site-directed mutagenesis that conserved hydrophilic residues from a carbohydrate-binding site of human galectin-1. *Glycoconj. J.* **11**, 437-442
44. Lobsanov, Y.D., Gitt, M.A., Leffler, H., Barondes, S.H., and Rini, J.M. (1993) X-ray crystal structure of the human dimeric S-Lac lectin, L-14-II, in complex with lactose at 2.9-Å resolution. *J. Biol. Chem.* **268**, 27034-27038
45. Liao, D.-I., Kapadia, G., Ahmed, H., Vasta, G.R., and Herzberg, O. (1994) Structure of S-lectin, a developmentally regulated vertebrate β-galactoside-binding protein. *Proc. Natl. Acad. Sci. USA* **91**, 1428-1432
46. Bourne, Y., Bolgiano, B., Liao, D.-I., Strecker, G., Cantau, P., Herzberg, O., Feizi, T., and Cambillau, C. (1994) Crosslinking of mammalian lectin (galectin-1) by complex biantennary saccharides. *Nature Struct. Biol.* **1**, 863-870
47. Hirabayashi, J., Ubukata, T., and Kasai, K. (1996) Purification and molecular characterization of a novel 16-kDa galectin from the nematode *Caenorhabditis elegans*. *J. Biol. Chem.* **271**, 2497-2505
48. Arata, Y., Akimoto,Y., Hirabayashi, J., Kasai, K., and Hirano, H. (1996) An immunohistochemical study of the 32-kDa galectin (β-galactoside-binding lectin) in the nematode *Caenorhabditis elegans*. *Histochem. J.* **28**, 201-207

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012