Glycoconjugate Journal (1994) **11**:437–442

Further evidence by site-directed mutagenesis that conserved hydrophilic residues form a carbohydrate-binding site of human galectin-1

JUN HIRABAYASHI* and KEN-ICHI KASAI

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

Received 19 April 1994, revised 7 June 1994

To identify critical amino acid residues for carbohydrate binding of galectins (soluble β-galactoside-binding lectins found in the animal kingdom), site-directed mutagenesis was performed on human galectin-1. On the basis of the previous results (Hirabayashi and Kasai (1992) *J Biol Chem* **266**:23648–53), more systematic mutagenesis experiments were performed in order to confirm the concept that conserved hydrophilic residues play a central role. When a homologous substitution was made for highly conserved His44, Arg48 or Asn61, the resultant mutant (H44Q, R48H or N61D, respectively) almost completely lacked carbohydrate-binding ability, as found previously for Asn46, Glu71 and Arg73 mutants. This suggests these six hydrophilic residues are essential. On the other hand, when less conserved Lys63, Arg111 or Asp125 were substituted, the resultant mutant (K63H, R111H or D125E, respectively) retained almost the same affinities to asialofetuin and lactose as the wild-type galectin. Therefore, none of these residues is directly involved in the binding. These results, together with the previous observation that the above six essential residues are all encoded in the largest exon of the gene and are located close to each other in the central, most hydrophilic region of the protein, suggest that the residues form a carbohydrate-binding site of galectin.

Keywords: galectin; mutagenesis; carbohydrate-recognition domain

Abbreviations: EDTA-PBS, 2 mM EDTA, 20 mM Na-phosphate, pH 7.2, 150 mM NaCl; MEPBS, EDTA-PBS containing 4 mM β-mercaptoethanol; IPTG, isopropyl-β-(D)-thiogalactoside; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Introduction

Soluble β-galactoside-binding lectin was first identified in the teleost *Electricus electriticus* [1], but subsequently a number of proteins having similar properties were found in various types of vertebrate cells and tissues including embryonic and tumourigenic ones [2]. For them, a new systematic name 'galectin' (meaning *gal*actose-specific *lectin*) has recently been proposed [3]. The present authors recently discovered the first invertebrate galectin, a novel type of galectin from a very primitive metazoan, the nematode *Caenorhabditis elegans* [4]. The nematode galectin (M_r 32 kDa) is composed of two tandemly repeated domains (designated tandem-repeat type), each of them being homologous to vertebrate 14 kDa galectins (approximately 30% amino acid identities). Only lactose is required for extraction of the nematode galectin. Like vertebrate galectins the nematode galectin does not require a metal–ion for the saccharide-binding activity. A cDNA for such a tandem-repeat type galectin has also been cloned from rat intestine [5]. Moreover, two related galectin cDNAs (corresponding to proteins of approximate M_r 14 kDa) were cloned from a much more primitive multicellular animal, the marine sponge *Geodia cydonium* [6]. The existence of various types of galectin molecules in a wide range of animal species implies their fundamental importance in cell–cell or cell–substratum recognition. Though their true biological function(s) have not been fully established, it is strongly suggested that the β-galactoside recognition process is basic to almost all multicellular animals. In this regard, a molecular level approach is needed to understand not only structure–function relationships but also evolutionary meaning.

We have already approached this goal by site-directed mutagenesis [7] of human galectin-1 (14 kDa lectin), one of the best studied galectins originally purified from the placenta [8–10], and later shown to occur in the lung [11],

* To whom correspondence should be addressed.

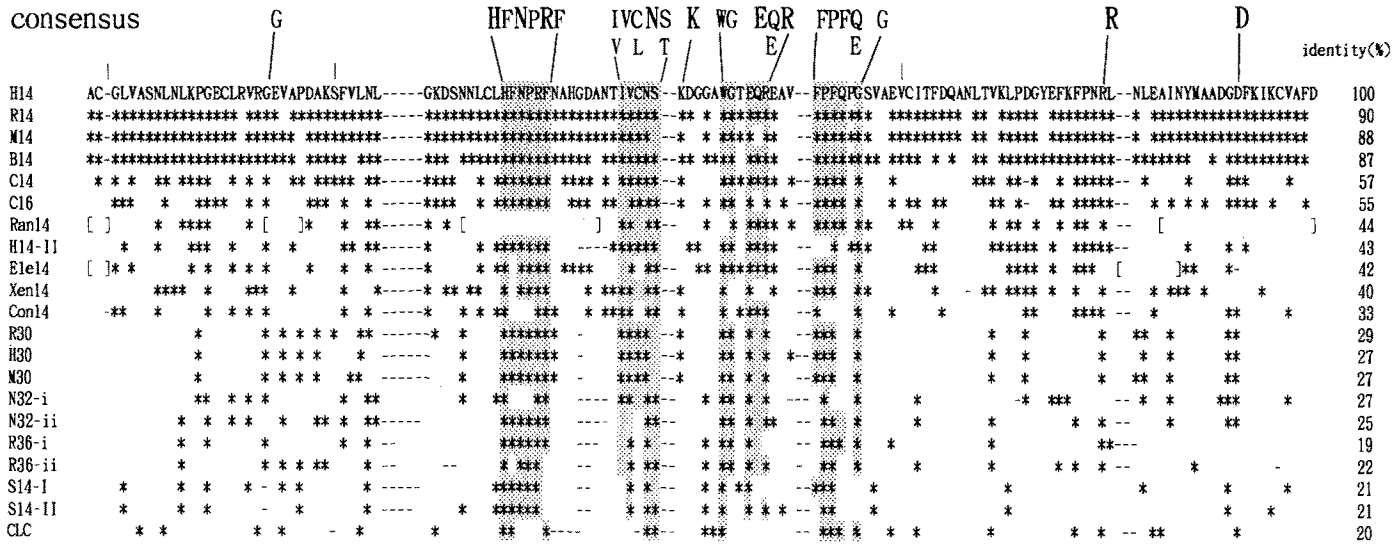

Figure 1. Alignment of amino acid sequences of metazoan galectins. Only positions of residues which are identical to those of human galectin-1 (H14, top line) are shown by asterisks for simplicity. Strongly conserved residues are also shown at the top. Dashes and open spaces represent deletions introduced for maximum homology and different residues, respectively. Regions for which the sequence has not so far been determined are enclosed by brackets. Percentage identities to human galectin-1 are indicated on the right. Deletions are not included in the calculation. Abbreviations; R14, rat galectin-1 (14 kDa); M14, mouse galectin-1 (14 kDa); B14, bovine galectin-1 (14 kDa); C14, chicken galectin having an apparent molecular weight of 14 kDa; C16, chicken galectin having an apparent molecular weight of 16 kDa; Xen14, *Xenopus* skin galectin (14 kDa); Ele14, electric eel galectin (14 kDa); Con14, conger eel galectin (14 kDa); Ran14, Rana oocyte galectin (14 kDa); R30, rat galectin-3 (30 kDa); M30, mouse galectin-3 (30 kDa); H30, human galectin-3 (30 kDa); N32-i and N32-ii, nematode galectin (32 kDa) domains 1 and 2, rspectively; R32-i and R-32-ii, rat galectin-4 (36 kDa) domains 1 and 2, respectively; S14-I and S14-II, two sponge isolectins (14 kDa). Charcot-Leyden crystal (CLC) protein, for which no carbohydrate-binding activity has been reported, is also aligned for comparison. Large capitals denote the target amino acids, which were mutagenized in this study.

muscle [12], spleen [13], brain [14] and HL60 cells [15]. Previous results suggested the importance of conservative hydrophilic residues (Asn46, Glu71 and Arg73), because substitutions of any of them resulted in the apparent loss of the carbohydrate-binding activity, whereas none of the cysteine or tryptophan residues was essential. Though the latter finding was contrary to the widely accepted assumption, the importance of hydrophilic residues was supported by recent X-ray crystallographic analyses. Most of the conservative hydrophilic residues of the assumed binding sites were shown to form an extensive network of hydrogen bonds with a ligand saccharide molecule [16, 17]. Therefore, it is expected that these residues will be conserved, and if they are not, the carbohydrate specificity should be modified.

To date, 18 galectins have been sequenced, including 13 proto-type, three chimera-type, and two tandem-repeat type galectins (for review see [2]). Apparently, the most conserved residues are located in the central region (Fig. 1). In particular, six hydrophilic residues are well conserved; i.e. His44 (residue numbers are those of human galectin-1), Asn46, Arg48, Asn61, Glu71 and Arg73 (100%, 89%, 100%, 100%, 100% and 85%, respectively). It is noteworthy that perfectly conserved Arg48 and Glu71 are replaced by cysteine and glutamine, respectively, in Charcot-Leyden crystal protein, which has lysophospholipase activity and sequence homology with the galectin family, but shows no evidence of sugar-binding activity [18] (also aligned for comparison in Fig. 1). The conserved central region of galectins has been shown to be encoded by a single exon in the cases of both proto-type [19–22] and chimera-type [23] molecules. His44, Arg48 and Asn61, which had not been mutagenized in the previous study, were substituted with homologous amino acids. For comparison, we also mutagenized Arg111 and Asp125, both of them being well conserved (80% and 60%, respectively) but encoded by another exon corresponding to the C-terminal part, and Lys63, encoded by the same exon but much less conserved (55%).

Materials and methods

Mutagenesis

Site-directed mutagenesis was performed as described previously [7, 24] by using an Amersham *in vitro* site-directed mutagenesis kit (ver. 2) and 5′-phosphorylated mutagenic primers; 5′-CTG CAC TTC GAC CCT CGC-3′ (N46D: 18-mer; melting temperature (T_m), 56 °C), 5′-C AAC CCT CAC TTC AAC GC-3′ (R48H: 18-mer; T_m, 54 °C), 5′-C GTG TGC GAC AGC AAG G-3′ (N61E: 17-mer; T_m,

Table 1. Summary of mutagenesis.

Mutant	*Host for expression*	*Yield (mg per l culture)*	*Binding to asialofetuin*	*I_{50} of lactose (mM)*
Wild type	Y1090	2.00	Yes	0.54
H44Q	HB101	—	No	—
N46D*	TG1	—	No	—
R48H	HB101	—	No	—
N61D	HB101	—	No	—
K63H	HB101	0.69	Yes	0.48
W68Y*	HB101	1.66	Yes	1.1
E71Q*	TG1	—	No	—
R73H*	Y1090	—	No	—
R111H	HB101	2.35	Yes	0.48
D125E	HB101	1.36	Yes	0.36

* Denotes previously reported mutants [7].

52 °C), 5′-GC AAC AGC CAC GAC GGC GG-3′ (K63H: 19-mer; T_m, 58 °C), and 5′-C CCC AAC CAC CTC AAC C-3′ (R111H: 17-mer; T_m, 54 °C), 5′-GCT GAC GGT GAA TTC AAG ATC-3′ (D125E: 21-mer, T_m, 60 °C). Single-stranded M13tv18 DNA, which contains the anti-sense sequence of wild-type human galectin-1, was used as a template for mutagenesis. Replacement of target base(s) was confirmed by either direct dideoxy nucleotide sequencing or dot-blot hybridization by using a mutagenic primer as a probe (24).

For expression, the mutated lectin gene was cut by *Bam* HI digestion from the respective replicative form and then recloned into an expression vector, pUC540(Kan^R) [25]. Appropriate *E. coli* strains, typically *recA*-deficient HB101, were transformed. Clones expressing lectin antigens were chosen by colony immunostaining with a specific antiserum [26]. Culture of cells, induction and extraction of mutant lectins were conducted essentially as described previously [7, 26]. Active mutant lectins were purified by asialofetuin-agarose column chromatography.

Asialofetuin–Toyopearl retardation assay

In order to assess the binding ability of mutant proteins that have reduced activity, and which had not been firmly adsorbed on the asialofetuin agarose in the above experiment, retardation analysis was performed, based on high-performance affinity chromatography on asialofetuin–Toyopearl as described previously [7]. Briefly, *E. coli* lysate containing 1 mg of protein was applied to a column of asialofetuin–Toyopearl (4.6 × 50 mm; 9.7 mg ml^{-1} of gel), which had been equilibrated with EDTA-PBS (2 mM EDTA, 20 mM Na-phosphate, pH 7.2), at a flow rate of 0.5 ml min^{-1}. Protein was monitored by measuring fluorescence emitted from tryptophan (excitation 280 nm; emission 350 nm). Immediately after injection, the effluent was fractionated on a time basis (6 s per fraction). Portions of the fractions were subjected to Western-blotting analysis by using anti-human galectin-1 antiserum as a probe, and the peak fraction which contained the maximum amount of the lectin antigen was identified.

Lectin–asialofetuin binding assay

Purified active mutant lectins were subjected to lectin–asialofetuin binding assay to assess quantitatively their sugar-binding activity [9]. Briefly, purified mutant lectin (250 ng) was reacted with an asialofetuin-coated immunoplate (Nunc, Maxisorb) in the presence of lactose (0.01–10 mM), and the bound lectin was determined by a double-antibody method with anti-human galectin-1 antiserum and horseradish peroxidase-conjugated goat anti-rabbit IgG. TMBlue (Schleicher & Schull) was used as a substrate. Affinity was evaluated from a dose-response curve, in terms of lactose concentration required for 50% inhibition (I_{50}) of lectin binding to the immobilized asialofetuin under the conditions employed.

Results

Expression of mutant lectin genes

Base substitutions generated by mutagenesis were confirmed either by dot-blot hybridization using a corresponding mutagenic primer as a probe, or direct dideoxy nucleotide sequencing. A mutant lectin gene fragment was cut from the vector, inserted into an expression vector, pUC540(Kan^R), carrying an isopropyl-β-(D)-thiogalactoside (IPTG)-inducible *tac* promoter, and used for transformation of appropriate *E. coli* strains. *E. coli* colonies expressing the lectin antigen were chosen by colony immuno-blot analysis [26]. Obtained clones were propagated in LB medium containing ampicillin and kanamycin. Lectin expression was induced by adding IPTG to give a final concentration of 0.1 mM at 37 °C for 4 h, and the mutant lectin was extracted by sonication as described previously [7].

For purification, *E. coli* lysate from a 2-l culture was applied to an asialofetuin-agarose column (10 ml vol), and after extensive washing of the column with MEPBS (EDTA-PBS containing 4 mM β-mercaptoethanol), adsorbed protein was eluted with the same buffer supplemented with 0.1 M lactose. The lectin peak was detected in the cases of K63H, R111H and D125E, but not with the others; i.e. H44Q, R48H and N61D (data not shown). The former group of mutants proved to be pure in SDS-PAGE (Fig. 2). This suggests that the binding activity of the latter group of mutants was reduced significantly, while the former group retained activity. Both groups of mutants were subjected to the following assay for more detailed analysis.

Asialofetuin–Toyopearl retardation assay

The sugar-binding ability of the mutants with reduced activity (i.e. H44Q, R48H and N61D) was analysed

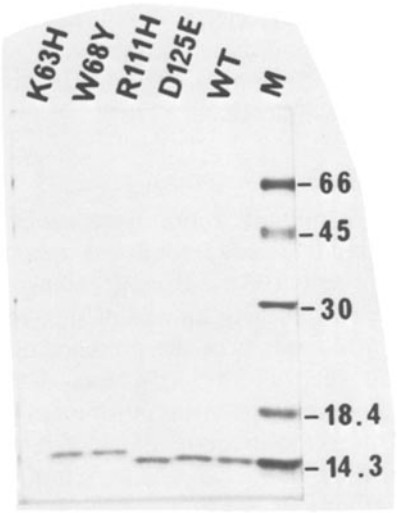

Figure 2. SDS-PAGE analysis of purified human galectin-1 mutants. Active mutant lectins purified on an asialofetuin-agarose column were analysed by SDS-PAGE using 14% gel in the presence of β-mercaptoethanol and stained with Coomassie Brilliant Blue. Abbreviations; K63H, Lys63 substituted with His; W68Y, Trp68 substituted with Tyr(7); R111H, Arg111 substituted with His; D125H, Asp125 substituted with Glu; WT, wild-type human galectin-1.

by means of a semi-quantitative method, asialofetuin–Toyopearl retardation assay, which is based on analytical affinity chromatography. *E. coli* cell lysates were applied to the column, and eluted at a relatively low flow rate to assure equilibrium between the stationary and moving phases. None of the inactive mutants showed a lectin peak after 0.1 M lactose elution. On the other hand, a clear fluorescent peak due to tryptophan was observed for all of the active mutants (K63H, R111H and D125E; data not shown) as well as the wild-type galectin-1 (Fig. 3a).

In order to determine the elution volume, both flow-through and retarded fractions were collected and analysed by Western-blotting in the cases of H44Q, R48H and N61D. In all cases examined, a specific lectin band (14 kDa) was detected in flow-through fractions, i.e. fractions 3 and 4 (corresponding to elution volumes of 0.75 ml and 1.0 ml, respectively; the column volume was 0.83 ml (Fig. 3b). With R48H, the lectin band in fraction 4 appears slightly more intense than that in fraction 3. However, the extent of retardation was found to be small for these mutants, and it was concluded that they retained no significant binding activity.

Lectin–asialofetuin-binding assay for active mutants

Mutant lectins which retained binding activity (K63H, R111H and D125E) were analysed more quantitatively. Their affinities to lactose were determined in terms of I_{50} in the lectin–asialofetuin-binding assay [7, 9], and were compared with that obtained for the wild-type recombinant human galectin-1. The I_{50} values for K63H, R111H and D125H were calculated as 0.40 mM, 0.48 mM and 0.36 mM, respectively (Fig. 4). The values are essentially the same as that obtained for the wild type, 0.54 mM, if we take the

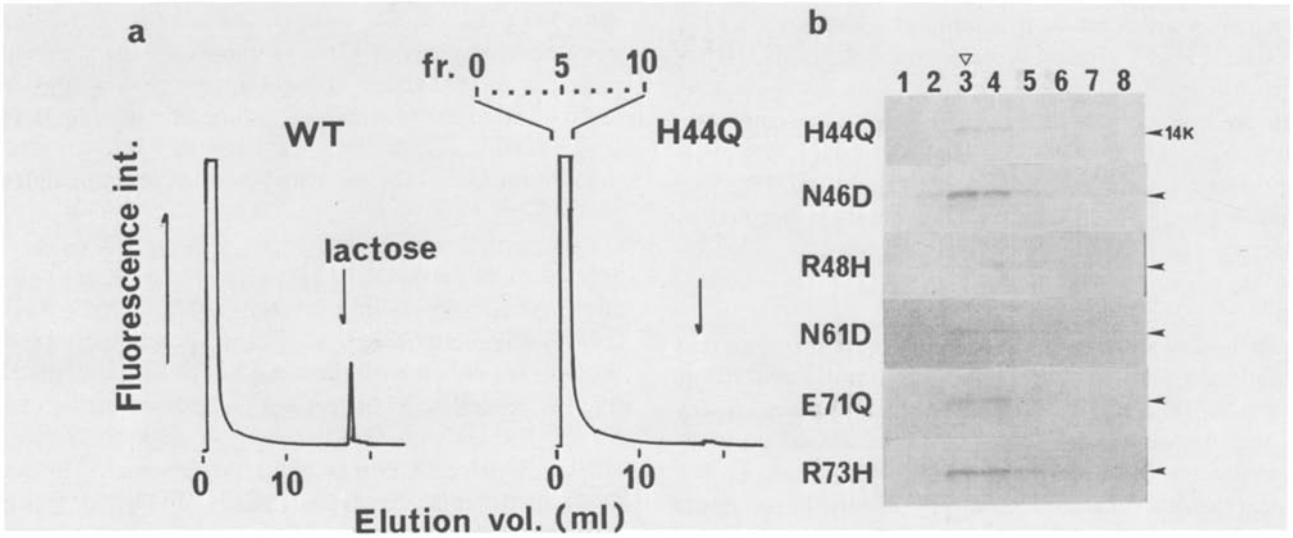

Figure 3. Asialofetuin–Toyopearl column retardation assay. (a) Examples of elution profiles of analytical asialofetuin–Toyopearl chromatography: *left*, wild type; *right*, H44Q mutant. Cell lysates each containing 1 mg of protein were applied to a column of asialofetuin–Toyopearl (4.6 × 50 mm; 9.7 mg per ml gel) equilibrated with EDTA-PBS at a flow rate of 0.5 ml min^{-1}. Protein elution was monitored by measuring tryptophan fluorescence emission (excitation and emission wavelengths, 280 nm and 350 nm, respectively). Bound lectin was eluted with EDTA-PBS containing 0.1 M lactose (arrow). For non-adsorbed mutant lectins, flow-through (fractions 2–4) and retarded fractions (fractions 5–10, 0.25 ml per fraction) were subjected to Western blotting analysis. (b) The result of Western blotting. An aliquot (0.02 ml) of each fraction obtained above was analysed by Western blotting using specific antiserum raised against human placenta lectin (galectin-1). The calculated flow-through position (0.83 ml) is indicated by an open triangle.

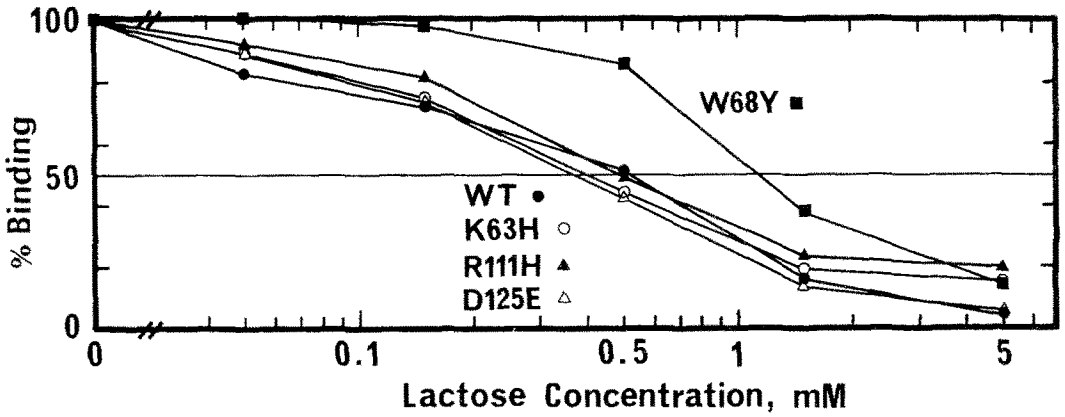

Figure 4. Results of lectin–asialofetuin binding assay. Sugar-binding ability of purified active mutant lectins, K63H, R111H and D125E was analysed in terms of lactose concentration inhibiting lectin binding to an asialofetuin-immobilized microtitre plate by 50% (I_{50}) as described under Materials and methods. The results for wild-type galectin-1 and W68Y reported in the previous paper [7] are also shown for comparison.

experimental error into consideration. This indicates that these mutants almost completely retained their binding activity for lactose. Figure 4 also shows that the previously described mutant W68Y retained significant but weaker affinity than the wild-type lectin. This again confirmed that Trp68 is not essential for the binding activity, but does enhance the interaction.

Discussion

In the previous mutagenesis experiments, attention had been paid mainly to cysteine and tryptophan residues [7], because they had long been believed to be essential to galectin function. However, the experimental results indicated that some hydrophilic residues such as Asn46, Glu71 and Arg73 are more important. In the present study, further experiments focusing on this point were carried out, and both well conserved and less well conserved hydrophilic residues were substituted. The results strongly suggested that His44, Arg48 and Asn61, as well as Asn46, Glu71 and Arg73, are critical for carbohydrate binding. Four (His44, Arg48, Asn61 and Glu71) of these six critical residues are conserved perfectly and two (Asn46 and Arg73) are conserved nearly perfectly with a few exceptions for non-mammalian galectins. In Charcot-Leyden eosinophilic crystal protein, four of them (Asn46, Arg48, Glu71 and Arg73) are replaced. This may explain why this protein does not show carbohydrate-binding activity.

On the other hand, homologous substitutions of Lys63 (only 55% conserved among galectins), Arg111(80% conserved) or Asp125 (60% conserved) did not affect the affinity to lactose as far as examined by lectin–asialofetuin-binding assay. They are evidently not involved in the binding function, though the latter two (Arg111 and Asp125) may be important for dimerization, because the C-terminal part of galectin has been suggested to be involved in dimer formation [27, 28].

The present results strongly suggest involvement of the six hydrophilic residues in the galectin-carbohydrate interaction, which is possibly achieved by extensive hydrogen bonding. The importance of Asn46 is also supported by the fact that the first domain of the nematode tandem-repeat type lectin, in which the corresponding residue is replaced with serine, has weaker carbohydrate-binding activity than the second domain (Arata, Hirabayashi, and Kasai, unpublished result).

During the preparation of this manuscript, Lobsanov *et al.* reported the three-dimensional structure, based on X-ray crystallography, of human galectin-2 (43% identical to human galectin-1) which formed a complex with a lactose molecule [29]. Their result is completely consistent with our present result: they proposed hydrogen bonds between galactose 4-OH and His45 (acceptor), Asn47 (acceptor) and Arg49 (donor) (corresponding to His44, Asn46 and Arg48, respectively of galectin-1), between galactose 6-OH and Asn58 (donor) and Glu69 (acceptor) (corresponding to Asn61 and Glu71, respectively), and between glucose 3-OH and Glu68 (acceptor) and Arg70 (acceptor) (corresponding to Glu71 and Arg73, respectively). The importance of Van der Waals interaction between the lactose molecule and conservative hydrophobic residues, Val56 and Trp69 (residue numbers are those of galectin-2), was also suggested. Though the unimportance of the tryptophan had been proved by the mutagenesis experiment in galectin-1 [7], we have not yet substituted the invariant valine.

Galectins seem to have conserved specificity for β-galactosides throughout metazoan evolution of almost 100 million years, though their biological meaning is not yet fully understood. From the molecular-evolutionary viewpoint, the recent finding that the topology of galectin resembles that of legume lectins [29] may imply that both animal and plant lectins evolved from an ancestral polypeptide, which may have had a weak carbohydrate-binding property. The conserved β-galactoside specificity in the

galectin family is remarkable in contrast to the diverse carbohydrate specificities of legume lectins and also to another group of animal lectins, i.e. C-type lectins [30]. It seems highly likely that galectins, together with their possible endogenous partners, lactosaminoglycans have an as-yet-unidentified key biological role on the basis of 'galactose recognition', as has been discussed recently by one of the present authors [31].

Acknowledgements

This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Sciences and Culture of Japan, and by grants from the Naito Foundation, and Kihara Memorial Yokohama Foundation for the Advancement of Life Sciences.

References

1. Teichberg VI, Silman I, Beitsh DD, Resheff G (1975) *Proc Natl Acad Sci USA* **72**:1383–7.
2. Hirabayashi J, Kasai K (1993) *Glycobiology* **3**:297–304.
3. Barondes SH, Castronovo V, Cooper DNW, Cummings RD, Drickamer K, Feizi T, Gitt MA, Hirabayashi J, Hughes RC, Kasai K, Leffler H, Liu F-T, Lotan R, Mercurio AM, Monsigny M, Pilli S, Poirer F, Raz A, Rigby PWJ, Rini JM, Wang JL (1994) *Cell* **76**:597–8.
4. Hirabayashi J, Satoh M, Kasai K (1992) *J Biol Chem* **267**:15485–90.
5. Oda Y, Herrmann J, Gitt M, Turck CW, Burlingame AL, Barondes SH, Leffler H (1993) *J Biol Chem* **268**:5929–39.
6. Pfeifer K, Haasemann M, Gamulin V, Bretting H, Fahrenholz F, Muller WEG (1993) *Glycobiology* **3**:179–84.
7. Hirabayashi J, Kasai K (1991) *J Biol Chem* **266**:23648–53.
8. Hirabayashi J, Kasai K (1984) *Biochem Biophys Res Commun* **122**:938–44.
9. Hirabayashi J, Kawasaki H, Suzuki K, Kasai K (1987) *J Biochem* **101**:987–95.
10. Hirabayashi J, Kasai K (1988) *J Biochem* **104**:1–4.
11. Hirabayashi J, Ayaki H, Soma G, Kasai K (1989) *Biochim Biophys Acta* **1008**:85–91.
12. Abbott WM, Feizi T (1989) *Biochem J* **259**:291–4.
13. Sharma A, Chemelli R, Allen HJ (1990) *Biochemistry* **29**:5309–14.
14. Bladier D, Le Caer JP, Joubert R, Caron M, Rossier J (1991) *Neurochem Int* **18**:275–81.
15. Couroud PO, Cadesentini-Borzocz D, Bringman JS, Griffith J, McGrogan M, Nedwin GE (1989) *J Biol Chem* **264**:1310–16.
16. Quiocho FA (1986) *Annu Rev Biochem* **55**:287–315.
17. Weis WI, Drickamer K, Hendrickson WA (1992) *Nature* **360**:127–35.
18. Ackerman SJ, Corrette SE, Rosenberg HF, Bennett JC, Mastrianni DM, Nicholson-Weller A, Weller PF, Chin DT, Tennen DG (1993) *J Immunol* **150**:456–68.
19. Ohyama Y, Kasai K (1988) *J Biochem* **104**:173–7.
20. Gitt MA, Barondes SH (1991) *Biochemistry* **30**:82–9.
21. Gitt MA, Massa SM, Leffler H, Barondes SH (1992) *J Biol Chem* **267**:10601–6.
22. Chiariotti L, Wells V, Bruni CB, Mallucci L (1991) *Biochim Biophys Acta* **1089**:54–60.
23. Gritzmacher CA, Mehl VS, Liu F-T (1992) *Biochemistry* **31**:9533–8.
24. Hirabayashi J, Kasai K (1993) In *Lectins and Glycobiology* (Gabius H-J, Gabius S eds) pp. 482–91. Berlin: Springer-Verlag.
25. Hirabayashi J, Ayaki H, Soma G, Kasai K (1989) *FEBS Lett* **250**:161–5.
26. Hirabayashi J, Sakakura Y, Kasai K (1993) In *Lectins and Glycobiology* (Gabius, H-J, Gabius S eds) pp. 474–81. Berlin: Springer-Verlag.
27. Harrison FL, Wilson TJG (1992) *J Cell Sci* **101**:635–46.
28. Hirabayashi J, Kasai K (1992) *J Chromatogr* **597**:181–7.
29. Lobsanov YD, Gitt MA, Leffler H, Barondes SH, Rini JM (1993) *J Biol Chem* **268**:27034–8.
30. Hirabayashi J (1993) *Trends Glycosci Glycotechnol* **5**:251–70.
31. Hirabayashi J (1994) In *Lectins, Biology, Biochemistry, Clinical Biochemistry*, Vol. 10 (Bøg-Hansen TC ed.) Sigma (in press).