

RESEARCH ARTICLE

# Structural Basis Underlying the Binding Preference of Human Galectins-1, -3 and -7 for Gal $\beta$ 1-3/4GlcNAc

Tung-Ju Hsieh<sup>1</sup>, Hsien-Ya Lin<sup>1,3</sup>, Zhijay Tu<sup>1</sup>, Bo-Shun Huang<sup>1,3</sup>, Shang-Chuen Wu<sup>1,4</sup>, Chun-Hung Lin<sup>1,2,3,4</sup>\*

**1** Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan, **2** The Genomics Research Center, Academia Sinica, Taipei, Taiwan, **3** Department of Chemistry, National Taiwan University, Taipei, Taiwan, **4** Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan

\* [chunhung@gate.sinica.edu.tw](mailto:chunhung@gate.sinica.edu.tw)



OPEN ACCESS

**Citation:** Hsieh T-J, Lin H-Y, Tu Z, Huang B-S, Wu S-C, Lin C-H (2015) Structural Basis Underlying the Binding Preference of Human Galectins-1, -3 and -7 for Gal $\beta$ 1-3/4GlcNAc. PLoS ONE 10(5): e0125946. doi:10.1371/journal.pone.0125946

**Academic Editor:** Yves St-Pierre, INRS, CANADA

**Received:** January 28, 2015

**Accepted:** March 26, 2015

**Published:** May 6, 2015

**Copyright:** © 2015 Hsieh et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The authors confirm that all data underlying the findings are fully available without restriction. Data is available from the Protein Data Bank using the accession codes 4XBL (hGal1/LN1), 4XBN (hGal3-CRD/LN1), and 4XBQ (hGal7/LN1).

**Funding:** This work was supported by Academia Sinica, and Ministry of Science and Technology grants: AS-022316 (to C.-H. Lin), NSC-103-2113-M-001-023-MY3 (to C.-H. Lin), NSC-102-2113-M-001-001-MY3 (to C.-H. Lin), and NSC-102-2923-M-001-001-MY3 (to C.-H. Lin). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Galectins represent  $\beta$ -galactoside-binding proteins and are known to bind Gal $\beta$ 1-3/4GlcNAc disaccharides (abbreviated as LN1 and LN2, respectively). Despite high sequence and structural homology shared by the carbohydrate recognition domain (CRD) of all galectin members, how each galectin displays different sugar-binding specificity still remains ambiguous. Herein we provided the first structural evidence of human galectins-1, 3-CRD and 7 in complex with LN1. Galectins-1 and 3 were shown to have higher affinity for LN2 than for LN1, while galectin-7 displayed the reversed specificity. In comparison with the previous LN2-complexed structures, the results indicated that the average glycosidic torsion angle of galectin-bound LN1 ( $\psi^{LN1} \approx 135^\circ$ ) was significantly differed from that of galectin-bound LN2 ( $\psi^{LN2} \approx -108^\circ$ ), i.e. the GlcNAc moiety adopted a different orientation to maintain essential interactions. Furthermore, we also identified an Arg-Asp/Glu-Glu-Arg salt-bridge network and the corresponding loop (to position the second Asp/Glu residue) critical for the LN1/2-binding preference.

## Introduction

Galectins,  $\beta$ -galactoside-binding proteins, are characteristic of having one or two conserved carbohydrate recognition domains (CRDs) [1, 2]. Members of this family have been shown to participate in diverse biological functions, such as cell adhesion, cell growth regulation, and apoptosis via their interactions with  $\beta$ -galactoside-containing structures on cell surface, e.g., N-, O-linked glycoproteins, proteoglycans or glycolipids [3, 4]. More importantly, human galectins act as regulatory factors in many types of cancers by either inhibiting or promoting tumor growth [5]. Therefore, to identify selective ligands for human galectins provides not only a useful tool for dissecting how each galectin member interacts with specific glycan structures in correlation with cancer progression, but also a possible solution for the development of clinical therapeutics.

**Competing Interests:** The authors have declared that no competing interests exist.

All galectins are able to recognize Gal $\beta$ 1-3/4GlcNAc disaccharides, namely type 1 and 2 LacNAc (abbreviated as LN1 and LN2, respectively) that appear in a myriad of glycoconjugates. For instance, LN1 and LN2 are shown as the repeating structures in the non-reducing termini of lacto-series glycans, including blood group antigens. LN2 are constitutively expressed in all mammalian cell types, while LN1 are more tissue-specifically distributed mostly restricted to the epithelia of gastrointestinal and reproductive tract in humans [6]. Interestingly, the presence of LN1 but not the LN2 was recently found to be in association with the pluripotency of induced pluripotent stem cells (iPSCs) [7, 8]. Furthermore, studies on the chemical structures of the milk oligosaccharides produced by various mammalian species revealed that LN1-containing oligosaccharides predominate over LN2-containing in human milk [9]. It has been hypothesized to be a selective advantage for human that the acquisition of predominantly LN1-containing oligosaccharides may promote the growth of specific anti-pathogenic bifidobacteria in the infant colon and thus aid their survival [9].

Galectins share major structural homology in their CRDs, but different galectin members were shown to display deviated binding preference for LN1- and LN2-containing glycans [10–13]. Albeit several crystal structures of galectin/LN2 complexes are available, there is no report regarding to the structure of galectin/LN1 complex. Therefore, the molecular basis still remains enigmatic to underlie the distinct binding preferences of galectins for LN1/2.

To bridge this gap, we herein report the crystal structures including full-length human galectin-1 (hGal1), galectin-7 (hGal7) and the C-terminal CRD domain of galectin-3 (hGal3-CRD) in complex with LN1. In comparison with the analogous LN2-complexed structures (deposited in Protein Data Bank with ID codes of 1W6P [14], 1KJL [15], 5GAL [16] for hGal1, 3-CRD and 7, respectively), the results pinpoint that in the bound LN1 and LN2, N-acetylglucosamine moiety adopts a different orientation while the galactose keeps the same posture. Additionally, we identified a unique Arg-Asp/Glu-Glu-Arg salt-bridges network (Arg48<sup>hGal1</sup>-Asp54<sup>hGal1</sup>-Glu71<sup>hGal1</sup>Arg73<sup>hGal1</sup>, Arg162<sup>hGal3</sup>-Glu165<sup>hGal3</sup>-Glu184<sup>hGal3</sup>-Arg186<sup>hGal3</sup> and Arg53<sup>hGal7</sup>-Glu58<sup>hGal7</sup>-Glu72<sup>hGal7</sup>-Arg74<sup>hGal7</sup>) to be essential for the LN1- or LN2-binding preferences. The loop harboring Asp54<sup>hGal1</sup>, Glu165<sup>hGal3</sup> or Glu58<sup>hGal7</sup> was found to affect how the salt-bridge is arranged and associated with possible water-mediated interactions.

## Materials and Methods

### Protein preparation

Using a standard PCR-based cloning strategy, the coding region of the full-length hGal1 (residues 1–135), 3 (residues 1–250), 7 (residues 1–136) and CRD domain of hGal3 (residues 113–250) were generated and inserted into modified pET-15b (hGal1, hGal3 and hGal3-CRD) or pET-28a (hGal7) vector (Novagen) with in-frame N- and C-terminal 6xHis-tag, respectively. QuikChange mutagenesis method (Agilent Technologies) was applied to replace the corresponding codons of residues Glu165 and Arg186 in full-length hGal3 with Ala codon to allow the expression of mutant proteins hGal3-E165A and hGal3-R186A. All the wild type and mutated proteins were produced in *Escherichia coli* BL21 (DE3), with 0.5 mM IPTG induction for 16 h at 20 °C. Recombinant proteins were purified by Ni<sup>2+</sup>-affinity and size-exclusion chromatography to homogeneity. Purified proteins were then stored in gel-filtration buffer (25 mM Tris-HCl pH8.0, 300 mM NaCl and 5 mM  $\beta$ -mercaptoethanol) and concentrated to ~ 4, 20 and 9 mg/ml as determined by the method of Bradford for both Biolayer interferometry experiments and crystallization trials, respectively.

## Bilayer interferometry

LN1 and LN2 binding affinity of hGal1, hGal3 and hGal7 were quantitatively measured in 96-well microplates at 27 °C by Octet Red system (FortéBio). Specifically, biotinylated galectins were prepared according to the standard protocol (provided by FortéBio), and adjusted to final concentration of 1 μM in assay buffer condition (50 mM Tris-HCl pH 7.5, 300 mM NaCl and 5 mM β-mercaptoethanol). Biotinylated galectins were then immobilized on Super Streptavidin Biosensors (FortéBio, Inc.), while free streptavidin sites on biosensor were then blocked by incubation with biocytin (10 mg/ml) to avoid non-specific interactions. The assay was carried out by placing galectin-coated biosensors into the wells with a concentration series of 3-fold diluted LN1/LN2 solutions (200 μl per well, from a top concentration of 3 mM for hGal1 and hGal7, and 1 mM for hGal3) and measuring changes in layer thickness (in nanometers) of biosensors with time. Measurements were followed by 120 sec baseline step, 60 sec association step and 200 sec dissociation step. Baseline and dissociation steps were carried out in assay buffer only. All the data were processed and calculated using Fortebio software and the steady-state  $K_d$  values were derived from equilibrium responses (S1 Fig) and summarized in Table 1. Two additional parallel Super Streptavidin biosensors were coated with biotinylated galectin and biocytin, separately, and only incubated with assay buffer as double reference controls.

## Fluorescence polarization (FP)-based competition binding assay

The measurement of FP-based assay is based on the rotation speed of a fluorophore-containing compound bound to the protein counterpart (e.g. galectins in this study). The fluorophore rotates at a slower rate than when it is unbound, and the resulting fluorescence polarization is higher. In this study we carried out all measurements according to reported procedures [17, 18]. To the final sample volume (70 μl) in each assay was added a synthetic FITC-conjugated type 2 LacNAc (LN2-FITC; as the fluorescent probe or reference compound) to a final concentration of 0.1 μM. All the measurements were conducted in Tris buffer (12.5 mM, pH 7.4) with 200 mM NaCl and 5 mM β-mercaptoethanol at 4 °C. Data plotting, nonlinear regression analysis, and curve construction was done by Prism 5.0 software (GraphPad, San Diego, CA).

For direct binding assay of LN2-FITC, the data (anisotropy, A vs. hGal 1, 3 and 7 concentration, respectively) were fitted to the formula  $A = A_0 + A_{max} \times [hGal] / (K_d + [hGal])$  to estimate  $K_d^{Probe}$  value (S2 Fig), where  $A_0$  is the anisotropy value measured in the absence of hGal and  $A_{max}$  is the maximum value approached with increasing [hGal].

**Table 1. Dissociation constants ( $K_d$  in μM) of human galectins-1, 3, 7 and mutants for Galβ1-3/4GlcNAc disaccharides<sup>1,2</sup>.**

Protein	$K_d$ (μM)	
	Galβ1-3GlcNAc (LN1)	Galβ1-4GlcNAc (LN2)
hGal1	340 <sup>1</sup> / 6.74 ± 3.14 <sup>2</sup>	150 <sup>1</sup> / 3.59 ± 2.35 <sup>2</sup>
hGal3	93 <sup>1</sup> / 1.44 ± 0.28 <sup>2</sup>	33 <sup>1</sup> / 0.33 ± 0.064 <sup>2</sup>
hGal7	270 <sup>1</sup> / 6.63 ± 1.27 <sup>2</sup>	410 <sup>1</sup> / 19.69 ± 3.45 <sup>2</sup>
hGal3-E165A	230 <sup>1</sup>	280 <sup>1</sup>
hGal3-R186A	ND <sup>3</sup>	ND <sup>3</sup>

<sup>1</sup>The values are determined by Bilayer Interferometry at 300 K.

<sup>2</sup>The values are determined by fluorescence anisotropy at 277 K.

<sup>3</sup>The binding is too weak to be determined by Bilayer Interferometry.

doi:10.1371/journal.pone.0125946.t001

For FP-based competition assay, commercial LN1 (or LN2) molecules (Dextra, UK) at indicated concentration 300, 6 and 300  $\mu\text{M}$  were used as competitors to compete binding interaction between 0.1  $\mu\text{M}$  fluorescent probe (LN2-FITC) and selected concentration of hGal1 (120  $\mu\text{M}$ ), 3 (3  $\mu\text{M}$ ) and 7 (120  $\mu\text{M}$ ), respectively. The measured anisotropy value with competitor ( $A_{\text{competitor}}$ ) is used to calculate the amount of galectin-bound probe [PG] according to equation:

$$[\text{PG}] = [(A_{\text{competitor}} - A_0) / (A_{\text{max}} - A_0)] \times [\text{P}]_{\text{total}}$$

And therefore,  $K_d^{\text{Competitor}}$  value was further deduced by the following equations and summarized in [Table 1](#):

$$\begin{aligned} [\text{P}] &= [\text{P}]_{\text{total}} - [\text{PG}] \\ [\text{G}] &= K_d^{\text{Probe}} \times [\text{PG}] / [\text{P}] \\ [\text{CG}] &= [\text{G}]_{\text{total}} - [\text{PG}] - [\text{G}] \\ [\text{C}] &= [\text{C}]_{\text{total}} - [\text{CG}] \\ K_d^{\text{Competitor}} &= [\text{C}] \times [\text{G}] / [\text{CG}] \end{aligned}$$

[C], [P] and [G] are the concentrations of free competitor, probe, and galectins, respectively, and [CG] and [PG] are the concentrations of competitor-galectin complex and the probe-galectin complex, respectively.

## Crystallization and data collection

Crystals of recombinant hGal1, 3-CRD were grown at room temperature (298 K) using the hanging-drop vapor diffusion method from 2  $\mu\text{l}$  protein solution and 2  $\mu\text{l}$  reservoir solutions consisting of 0.1 M Tris pH 8.0, 0.2 M  $\text{Li}_2\text{SO}_4$ , 30% (w/v) PEG 3350 and of 0.1 M Tris pH 8.5, 0.2 M  $\text{MgCl}_2$ , 30% (w/v) PEG 4000, for hGal1 and hGal3-CRD respectively. The hGal1-LN1 and hGal3-CRD-LN1 complexes were obtained by soaking their native protein crystals with 20 mM type I N-acetyllactosamine (LN1) for more than a week. hGal7 protein solution at 9 mg/ml was incubated overnight with 2 mM LN1 molecule at 4  $^\circ\text{C}$ . Following overnight incubation, the protein-ligand mixtures were centrifuged at 13,000 rpm for 5 min to remove the precipitated protein. Cocrystals for the hGal7-LN1 complexes were carried out by the hanging-drop vapor diffusion method at room temperature by mixing 2  $\mu\text{l}$  protein solution and 2  $\mu\text{l}$  reservoir solution containing 0.1 M HEPES pH 7.5, 0.2 M  $\text{Li}_2\text{SO}_4$ , 0.1 M NaOAc and 25% (w/v) PEG 4000. The reservoir solutions supplemented with 10 to 20% glycerol were used for cryo-protection of the complex crystals. The crystals were then flash-frozen in liquid nitrogen and stored for synchrotron-radiation data collection. The diffraction data were processed using the HKL2000 program suite [19] with data statistics as summarized in [Table 2](#).

## Determination and refinement of the crystal structures

The crystal structures of all complexes were solved by molecular replacement with the PHENIX AutoMR [20] using previously published ligand-free galectin structures as the starting search models: PDB entries 1W6N [14], 2NMN [21] and 1BKZ [16] for hGal1, 3-CRD and 7, respectively. Model building was performed with PHENIX AutoBuild [20]. The resulting electron density maps were of good quality and show clearly the densities belonging to the bound LN1 molecules. The structure of LN1 molecule was created by JLigand version 1.0.35 in CCP4 software suite [22] and built into the density by using Coot [23]. Structures then underwent rounds

**Table 2. Data collection and refinement statistics.**

	hGal1-LN1	hGal3-CRD-LN1	hGal7-LN1
<b>Data collection</b>			
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	43.3, 58.2, 111.4	37.1, 57.6, 63.2	30.1, 55.4, 136.5
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	30–1.93 (2–1.93) <sup>1</sup>	30–2.208 (2.29–2.208)	30–2.23 (2.31–2.23)
<i>I</i> / <i>s</i>	29.67 (7.11)	38.35 (18.23)	14.59 (6.21)
Completeness (%)	99.5 (96.4)	99.47 (95.45)	98.67 (96.45)
Redundancy	6.7 (6.3)	7.7 (7.8)	4.8 (4.7)
<i>R</i> <sub>sym</sub> <sup>2</sup>	0.056 (0.246)	0.051 (0.132)	0.050 (0.277)
<b>Refinement</b>			
<i>R</i> <sub>work</sub> <sup>3</sup> / <i>R</i> <sub>free</sub> <sup>4</sup>	0.196 / 0.235	0.165 / 0.203	0.196 / 0.225
<i>B</i> -factors (mean)	34.40	32.00	30.00
r.m.s. deviations			
Bond lengths (Å)	0.010	0.008	0.007
Bond angles (°)	1.24	1.03	0.85
Ramachandran (%)			
Favored (%)	97	99	98
Outliers (%)	0	0	0
<b>PDB ID code</b>	4XBL	4XBN	4XBQ

<sup>1</sup>Statistics for data from the highest-resolution shell are shown in parentheses.

<sup>2</sup> $R_{sym} = (\sum \sum |I_{hkl} - \langle I \rangle|) / (\sum I_{hkl})$ , where the average intensity  $\langle I \rangle$  is taken overall symmetry equivalent measurements and  $I_{hkl}$  is the measured intensity for any given reflection.

<sup>3</sup> $R_{work} = (\sum ||F_o| - k|F_c||) / (\sum |F_o|)$ , where  $F_o$  and  $F_c$  are the observed and calculated structure factor amplitudes, respectively.

<sup>4</sup> $R_{free}$  was calculated for R factor using only an unrefined subset of reflections data (5%).

doi:10.1371/journal.pone.0125946.t002

of manual model rebuilding and refinement with Coot and PHENIX. Detailed refinement parameters are listed in [Table 2](#). The figures were generated in Pymol [24].

## Results and Discussion

### Binding affinity and preference of hGal1, hGal3 and hGal7 for Galβ 1-3/4GlcNAc

We first quantitated the LN1- and LN2-binding affinity of hGal1, 3 and 7 by two different methods, Biolayer interferometry [25] and FP-based competition assays [17, 18]. The resulting  $K_d$  values are summarized in [Table 1](#) to show the binding affinity of 3 human galectins with LN1 and LN2. Although the  $K_d$  values obtained by Biolayer interferometry ( $K_d^{BI}$ ) are generally about 2 orders of magnitude greater than those by FP-based competition assays ( $K_d^{FP}$ ), both measurements reveal a highly consistent trend. hGal7 showed significantly increased affinity for LN1 ( $K_d^{BI} = 270 \mu\text{M}$  or  $K_d^{FP} = 6.6 \mu\text{M}$ ) than for LN2 ( $K_d^{BI} = 410 \mu\text{M}$  or  $K_d^{FP} = 19.7 \mu\text{M}$ ), whereas hGal1 and hGal3 are more specific for LN2 ([Table 1](#)). Biolayer interferometry was known to measure the binding affinity and additional kinetic detail of the given compound/protein complex, such as kinetic constants of association and dissociation [26]. Recently Biolayer interferometry was applied to characterize the binding properties of galectins [27]. It is common that different methods measuring the same binding event often produce different  $K_d$

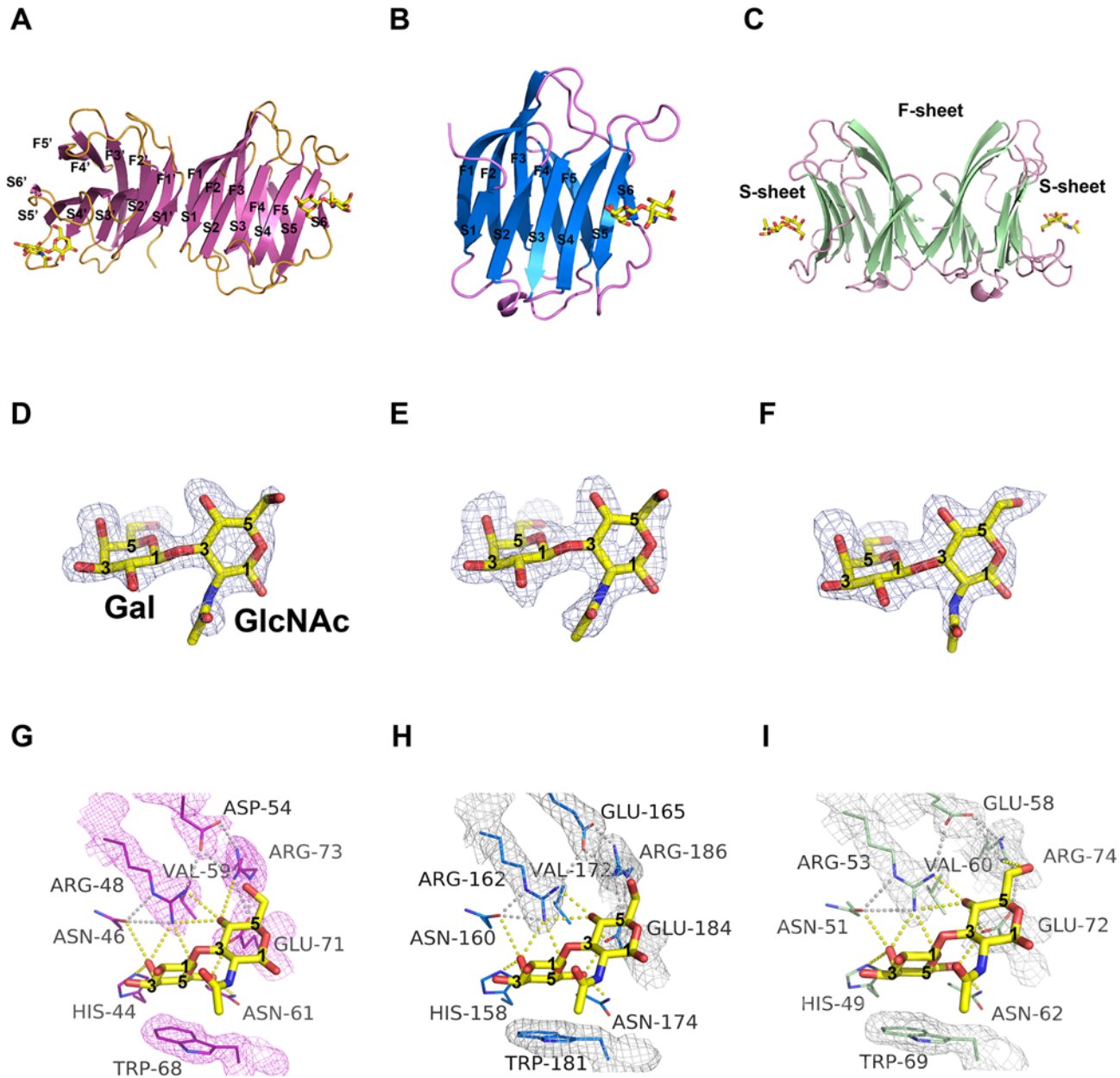
values. The difference could vary from 1 to 2 orders of magnitude [12, 28–30], but the most important issue is to see if there is a consistent trend on the binding when comparing different methods. Overall, the LN1/LN2 binding preference of galectins-1, -3 and -7 are not strict. For instance, only a maximum 3-fold difference in LN1 and LN2 binding affinity of hGal7 was observed in this study, suggesting the possibility of functional redundancy among members of galectin family [31, 32]. Given the feature of multivalent interactions in the context of galectin/glycan lattices, even 2- or 3-fold change in their individual interactions would result in substantially higher activity and/or a more dramatic effect [33]. As a matter of fact, several studies also correlate the binding preference of LN1 or LN2 with physiological activities [34]. Structural information is thus necessary to delineate the insight at molecular level.

### Overall structures of hGal1, 3-CRD and 7 in complex with LN1

hGal1, the CRD of hGal3 (hGal3-CRD) and hGal7 in complex with LN1 were crystallized by either a soaking or co-crystallization method. Their crystal structures were then determined between 1.9 and 2.2 Å by molecular replacement on the basis of the published ligand-free structures (PDB IDs: 1W6N [14], 2NMN [21] and 1BKZ [16] corresponding to hGal1, hGal3-CRD and hGal7, respectively) as the starting search models. Statistics of data processing and refinement parameters of the structures are summarized in Table 2. While only one monomer exists in the asymmetric unit of hGal3-CRD co-crystal, each asymmetric unit of hGal1 and hGal7 harbors a distinct symmetric dimer (Fig 1A–1C). Specifically, the dimer of hGal1 exists in a side-by-side manner, whereas hGal7 dimer is present in a back-to-back arrangement. Because the two protomers of the hGal1 and hGal7 dimers are almost identical to each other, we refer to chain A of each crystal structure in the following discussion. All the observed CRDs adopt a typical galectin fold which is composed of two antiparallel  $\beta$ -sheets of six (S1–S6) and five (F1–F5) strands, jointly forming a  $\beta$ -sheet sandwich structure and therefore named as S-sheets and F-sheets, respectively. The S1–S6  $\beta$ -strands constitute a concave surface to which  $\beta$ -galactoside-containing glycans are bound. Electron density belonging to the bound LN1 is clearly identified in the *F<sub>o</sub>-F<sub>c</sub>* electron density map (Fig 1D–1F), indicating that the LN1 molecule in the complex is well ordered and all sugar rings in the LN1 adopt a chair conformation. Of note, overall root mean square deviation (RMSD) values among the newly determined LN1-bound galectins (hGal1, hGal3-CRD and hGal7), the ligand-free and the LN2-bound galectins (PDB IDs: 1W6N, 3ZSM [35] and 1BKZ for ligand-free hGal1, hGal3-CRD and hGal7; PDB IDs: 1W6P [14], 1KJL [15] and 5GAL [16] for LN2-loaded hGal1, hGal3-CRD and hGal7) are quite small. The values of RMSD ( $C^\alpha$  atoms) between ligand-free and LN1-bound hGal1, 3 and 7 are of 0.57, 0.20 and 0.53 Å, respectively. RMSD values ( $C^\alpha$  atoms) of 0.10, 0.18 and 0.59 Å are presented between ligand-free and LN2-bound hGal1, hGal3 and hGal7 structures, respectively. Therefore, any differences due to the glycosidic linkages ( $\beta$ 1–3 or  $\beta$ 1–4) in LN1 and LN2 appear not to seriously distort the overall structure of hGal1, hGal3-CRD and hGal7.

Generally the galactose moiety (GAL) forms more H-bonds with the amino acid residues in the CRD than the N-acetylglucosamine moiety (GlcNAc), supporting the idea that GAL serves as a major recognition component. The GAL of LN1 interacts with a series of conserved residues located on S4–S6  $\beta$ -strands and the loop connecting S4 and S5 strands, which include His44<sup>hGal1</sup>/158<sup>hGal3</sup>/49<sup>hGal7</sup>, Asn46<sup>hGal1</sup>/160<sup>hGal3</sup>/51<sup>hGal1</sup>, Arg48<sup>hGal1</sup>/162<sup>hGal3</sup>/53<sup>hGal7</sup>, Asn61<sup>hGal1</sup>/174<sup>hGal3</sup>/62<sup>hGal7</sup> through hydrogen bond (H-bond) networks, and Trp68<sup>hGal1</sup>/181<sup>hGal3</sup>/69<sup>hGal7</sup> via van der Waals contacts (Fig 1G–1I). In particular, the conserved Arg48<sup>hGal1</sup>/162<sup>hGal3</sup>/53<sup>hGal7</sup> residues play an important role in mediating interactions between hGal1, hGal3, hGal7 and their corresponding LN1 molecules, respectively. Specifically, the Arg48<sup>hGal1</sup>/162<sup>hGal3</sup>/53<sup>hGal7</sup> residues not only bridge H-bonds to several oxygen atoms of LN1 including C4-OH, O5 of GAL and C4-OH





**Fig 1. Structural overview of hGal1, 3-CRD and 7 in complex with LN1.** (A-C) Ribbon representations of three LN1-hGal complexes where LN1 (shown in yellow stick model) is bound to hGal1 (A), hGal3-CRD (B) and hGal7 (C). Numbering of the  $\beta$ -strands of S-sheet (S1-S6) and F-sheet (F1-F5) is also shown as indicated. (D-F)  $F_o-F_c$  omit electron density map of LN1 (contoured at  $2.5\sigma$ ) bound to hGal1 (D), hGal3-CRD (E) and hGal7 (F). To make it clear, carbons 1, 3 and 5 of Gal and GlcNAc are labeled. (G-I)  $\beta$ -galactoside-recognition site of hGal1-LN1 (G), hGal3-CRD-LN1 (H) and hGal7-LN1 (I) complexes. Residues involved in LN1 recognition are highlighted with  $2F_o-F_c$  electron density (contoured at  $1\sigma$ ). Polar interactions among galectin residues and within galectin/LN1 complex are shown as gray and yellow dash lines, respectively.

doi:10.1371/journal.pone.0125946.g001

of GlcNAc, but also connect peripheral carbohydrate-interacting amino acid residues such as  $\text{Asn}46^{\text{hGal1}}/160^{\text{hGal3}}/51^{\text{hGal7}}$ ,  $\text{Asp}54^{\text{hGal1}}/\text{Glu}165^{\text{hGal3}}/58^{\text{hGal7}}$ , and  $\text{Arg}73^{\text{hGal1}}/186^{\text{hGal3}}/74^{\text{hGal7}}$  to form a characteristic interacting network of H-bonds and electrostatic interactions which are optimal for carbohydrate orientation in the binding curvature (Fig 1G–I) [16, 36]. The density maps of the LN1-interacting amino acid residues are shown in a satisfying quality to define the environment of the LN1-binding site.

## Structural basis for LN1- and LN2-binding preferences of hGal1, hGal3-CRD and hGal7

In accordance with the X-ray crystal structures, the numbers and distances of specific H-bond interactions involved in the recognition of LN1/ LN2 by hGal1, hGal3 and hGal7 are measured and summarized in [Table 3](#) as the basis to interpret their distinct binding specificity. Generally, hGal1 and hGal3 appear to have shorter distances (in average) in H-bonds to both GAL and GlcNAc moieties of LN2 than those of LN1. In contrast, hGal7 has more H-bonds to GAL moiety and a characteristic shorter distance with GlcNAc in LN1, as compared to those in LN2. These results correlate well with the aforementioned difference in the binding affinity. The static X-ray structures may not always correspond to the behavior of proteins in solution, further studies such as NMR or MD simulations based on these LN1/LN2-galectin complex structures would offer more insights with their binding dynamics in solution.

Overall, the LN1-recognition modes of hGal1, hGal3-CRD and hGal7 are quite similar to those observed in the LN2-bound complex structures. Structural superimpositions of LN2-complexed hGal1, hGal3-CRD and hGal7 structures (PDB IDs: 1W6P, 1KJL and 5GAL, respectively) with their LN1-loaded ones ([Fig 2A–2C](#)) indicate that the GAL moiety of LN1 and LN2 is overlapped well. Even though the GlcNAc moiety of LN1 and LN2 interacts with the same amino acid residues, the GlcNAc moiety was found to adopt a different orientation in LN1- and LN2-complex structures. Specifically, the average torsional angles  $\phi^{\text{LN1}}$  ( $-60^\circ$ , defined by  $\text{O5}^{\text{GAL}}-\text{C1}^{\text{GAL}}-\text{O3}^{\text{GlcNAc}}-\text{C3}^{\text{GlcNAc}}$  of LN1) and  $\phi^{\text{LN2}}$  ( $-66^\circ$ , defined by  $\text{O5}^{\text{GAL}}-\text{C1}^{\text{GAL}}-\text{O4}^{\text{GlcNAc}}-\text{C4}^{\text{GlcNAc}}$  of LN2) that depict the position of GAL relative to the glycosidic bond are quite similar ([Fig 2D](#)). On the other hand, the torsional angle  $\psi$  is to characterize the orientation of GlcNAc relative to the glycosidic bond. The average  $\psi^{\text{LN1}}$  ( $135^\circ$ , defined by  $\text{C1}^{\text{GAL}}-\text{O3}^{\text{GlcNAc}}-\text{C3}^{\text{GlcNAc}}-\text{C4}^{\text{GlcNAc}}$  of LN1) is dramatically different from the average  $\psi^{\text{LN2}}$  ( $-108^\circ$ , defined by  $\text{C1}^{\text{GAL}}-\text{O4}^{\text{GlcNAc}}-\text{C4}^{\text{GlcNAc}}-\text{C5}^{\text{GlcNAc}}$  of LN2). This  $\sim 240^\circ$  shift allows the three important OH groups of galectin-bound LN1 (C4-OH and C6-OH of Gal and C4-OH of GlcNAc) to form the binding interactions not only essential for the galectin recognition, but also homologous to those produced by the OH groups of galectin-bound LN2 (C4-OH and C6-OH of Gal and C3-OH of GlcNAc) ([Fig 2D](#)) [[10](#), [37](#), [38](#)]. The glycosidic torsional angles in this study were defined in previous reports [[16](#), [36](#)]. Notably the C4-OH group of GlcNAc in LN1 correlates with C3-OH group of GlcNAc in LN2 so that they are located at the equivalent position to form pivotal H-bonds with Arg48<sup>hGal1</sup>/162<sup>hGal3</sup>/53<sup>hGal7</sup> and Glu71<sup>hGal1</sup>/184<sup>hGal3</sup>/72<sup>hGal7</sup> that are highly conserved among human galectins ([Fig 2D](#)).

## Roles of Asp54<sup>hGal1</sup>/Glu165<sup>hGal3</sup>/58<sup>hGal7</sup> in mediating a unique salt bridge network

Furthermore, we found several subtle but major differences in the carbohydrate-binding sites of hGal1, hGal3-CRD and hGal7 ([Fig 2E–2G](#)). First, hGal1 and 7 have the electrostatic network consisting of Arg48<sup>hGal1</sup>/53<sup>hGal7</sup>, Asp54<sup>hGal1</sup>/Glu58<sup>hGal7</sup>, Glu71<sup>hGal1</sup>/72<sup>hGal7</sup> and Arg73<sup>hGal1</sup>/74<sup>hGal7</sup> [[16](#), [36](#)]. These residues adopt dissimilar conformations and orientations when interacting with LN1 or LN2 molecules ([Fig 2E and 2G](#)), while the corresponding network of hGal3-CRD (Arg162, Glu165, Glu184 and Arg186) interacts either with LN1 or LN2 in almost the same manner ([Fig 2F](#)). Neither of Asp54<sup>hGal1</sup>, Glu165<sup>hGal3</sup> and Glu58<sup>hGal7</sup> interacts with the disaccharides directly, but the way they interact with the Arg counterparts forming the salt bridge appears to be dissimilar. The unique salt-bridge network mediated by Glu58<sup>hGal7</sup> ([Fig 2G](#)) was different from those mediated by the corresponding residues Asp54<sup>hGal1</sup>, Glu165<sup>hGal3</sup> ([Fig 2E and 2F](#)). According to classification to define the salt bridge geometry [[39](#)], Glu58<sup>hGal7</sup> forms two weaker monodentate N-O bridges with Arg53<sup>hGal7</sup> and Arg74<sup>hGal7</sup> (the important



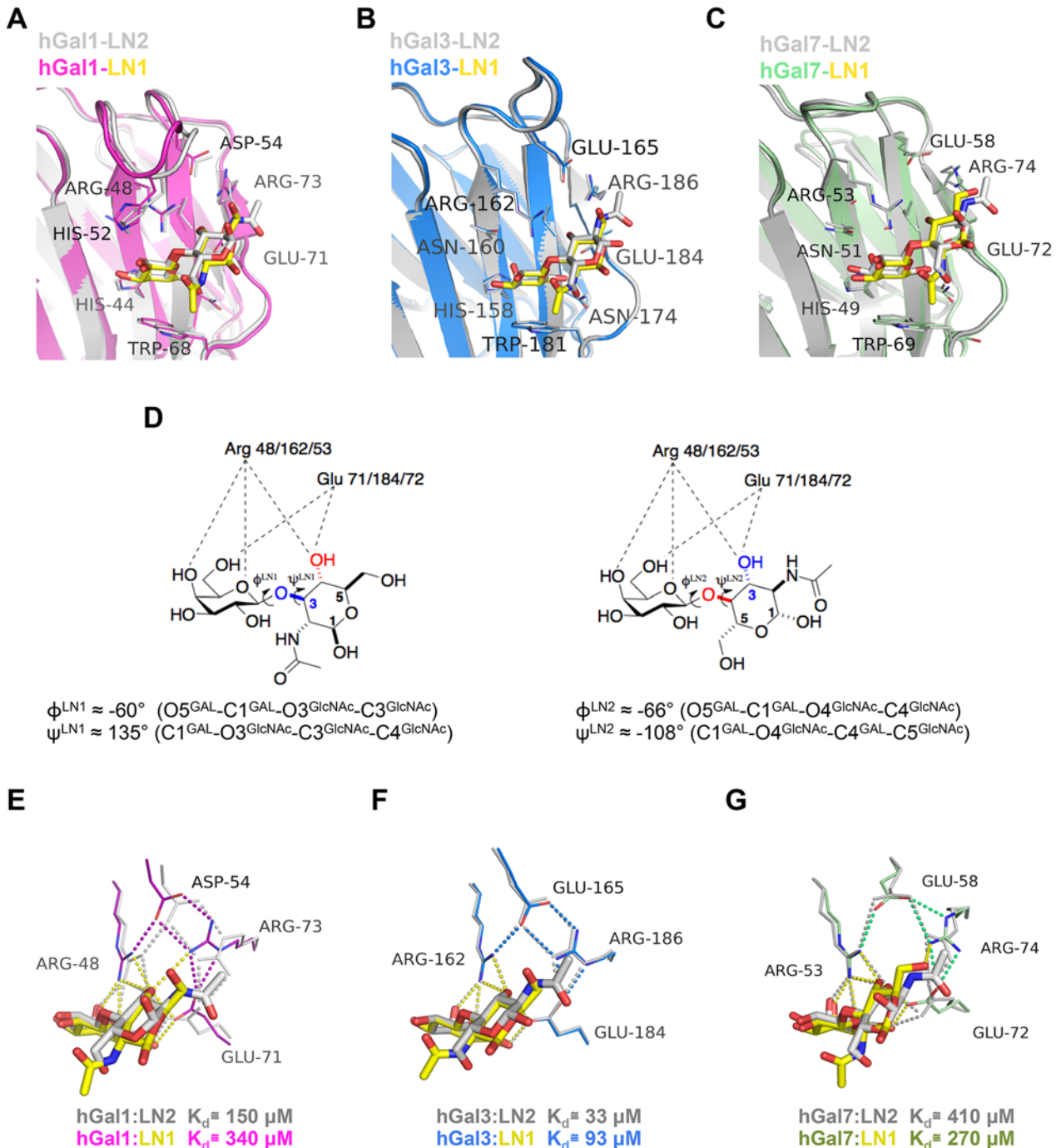
**Table 3. Comparison of polar interactions between human Galectins and LN1/2<sup>a</sup>.**

Sugar atom	Amino acid atom in hGal1/3/7	hGal1-LN1		hGal1-LN2 <sup>b</sup> (1W6P_Mol.A)	hGal3-CRD-LN1 Mol.A	hGal3CRD-LN2 <sup>b</sup> (1KJL)	hGal7-LN1		hGal7-LN2 <sup>b</sup> (5GAL_Mol.B)
		Mol.A	Mol.B				Mol.A	Mol.B	
GlcNAc O3	Glu 71/184/72 O <sub>ε</sub> 1	-	-	3.33	-	-	-	-	3.15
GlcNAc O3	Glu 71/184/72 O <sub>ε</sub> 2	-	-	2.59	-	2.54	-	-	2.62
GlcNAc O3	Arg 48/162/53 Nh1	-	-	2.73	-	2.77	-	-	3.07
GlcNAc O3	Arg 48/162/53 Nh2	-	-	2.99	-	3.15	-	-	3.42
GlcNAc O4	Glu 71/184/72 O <sub>ε</sub> 1	3.02	3.25	-	-	-	3.43	-	-
GlcNAc O4	Glu 71/184/72 O <sub>ε</sub> 2	2.65	2.40	-	2.58	-	2.48	2.87	-
GlcNAc O4	Arg 48/162/53 Nh1	2.88	2.90	-	2.78	-	2.77	2.54	-
GlcNAc O4	Arg 48/162/53 Nh2	3.22	3.21	-	3.08	-	2.92	3.07	-
GlcNAc O4	Arg 73/186/74 Nh2	3.31	3.47	-	-	-	-	-	-
GlcNAc N2	Water 23/01/-	-	-	2.80	-	2.9	-	-	-
Avg. interaction distance with GlcNAc		3.03		2.89	2.81	2.84	2.87		3.07
Gal O4	His 44/158/49 N <sub>ε</sub> 2	2.87	2.71	2.74	2.76	2.8	2.89	2.9	2.95
Gal O4	Asn 46/160/51 O <sub>ε</sub> 1	3.37	3.23	3.15	3.26	3.13	3.24	3.25	2.87
Gal O4	Arg 48/162/53 Nh2	2.77	3.02	3.01	2.93	2.86	2.91	2.72	2.6
Gal O5	Arg 48/162/53 Nh2	3.0	3.1	2.88	3.0	2.96	2.94	2.94	2.81
Gal O5	Glu 71/184/72 O <sub>ε</sub> 2	3.37	3.06	3.12	3.31	3.3	3.15	3.28	-
Gal O6	Asn 61/174/62 N <sub>ε</sub> 2	2.89	2.75	2.88	2.89	2.74	2.75	2.8	2.63
Gal O6	Glu 71/184/72 O <sub>ε</sub> 1	-	-	-	-	-	-	-	2.85
Gal O6	Glu 71/184/72 O <sub>ε</sub> 2	2.7	2.78	2.67	2.72	2.7	2.71	2.5	-
Avg. interaction distance with Gal		2.97		2.92	2.98	2.93	2.93		2.79

<sup>a</sup>hydrogen bond interactions were identified with program, Pymol. Values shown are distances in Å.

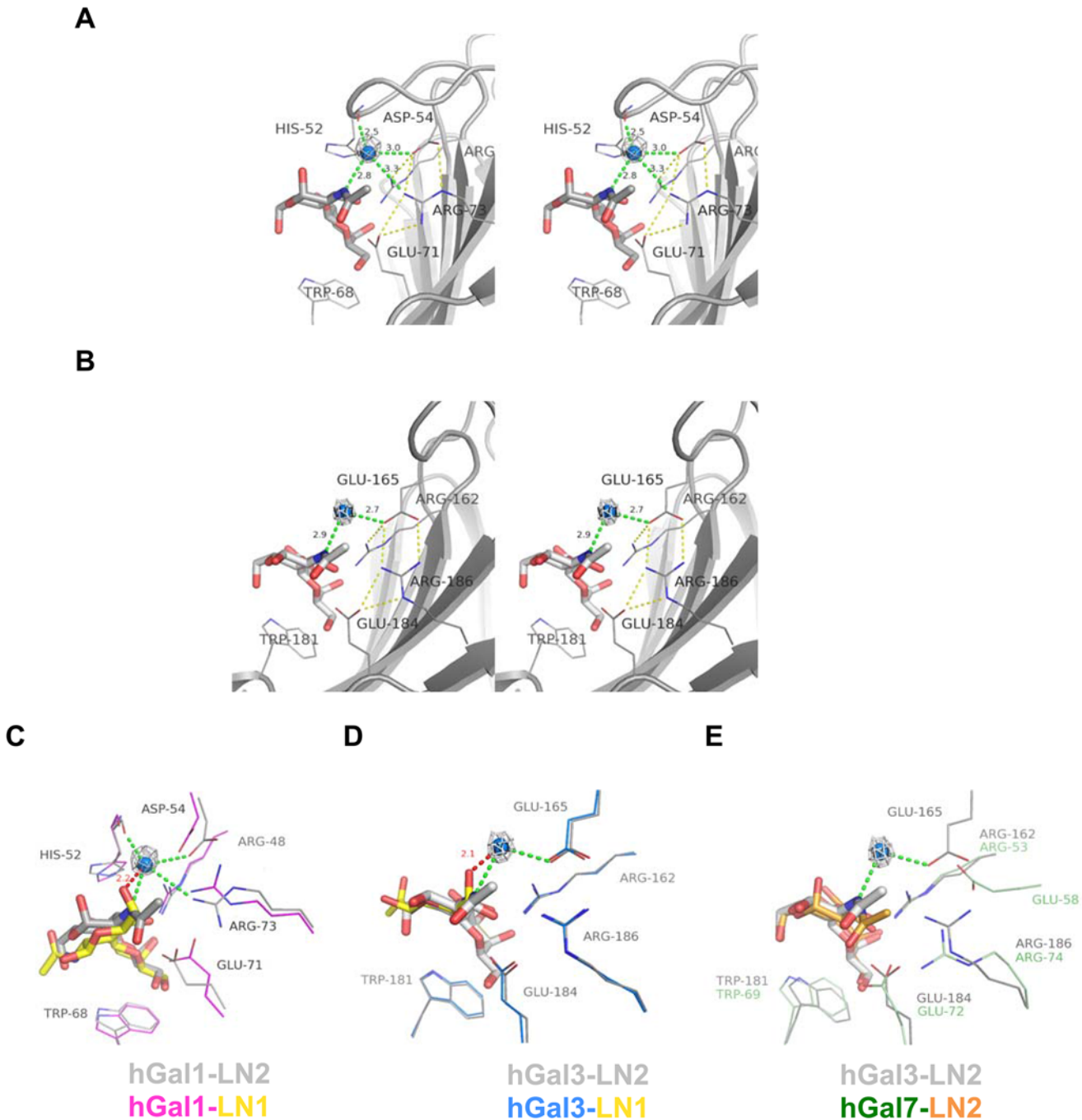
<sup>b</sup>LN2-complexed hGal1/3-CRD/7 structures reported in previous literatures were used for comparison.

doi:10.1371/journal.pone.0125946.t003



**Fig 2. Pairwise comparison of the  $\beta$ -galactoside-recognition site in the LN1-hGal1/3-CRD/7 and LN2-hGal1/3-CRD/7 complexes.** (A-C) Structures of hGal1 (PDB ID: 1W6P), hGal3-CRD (1KJL) and hGal7 (5GAL) in complex with LN2 (all shown in gray) were superimposed, respectively, with the LN1 (in yellow)-containing structures of hGal1 (pink), hGal3-CRD (cyan) and hGal7 (green). (D) Diagrams delineate the different interaction geometries of LN1 (left) and LN2 (right) with respect to hGal1/3-CRD/7. Definition and values of the glycosidic torsion angles ( $\phi$  and  $\psi$ ) for LN1 and LN2 molecules are also listed. (E-G) Close-up view of the unique salt bridge networks in hGal1 (E), hGal3-CRD (F) and hGal7 (G), with a superposition of their LN2 and LN1-complex structures. Polar interactions among LN2-complexed structures are shown in gray dash lines while the interactions among amino acid residues in LN1-loaded structures are indicated by colored dash lines according to hGal1 (pink), 3-CRD (blue) and 7 (green), respectively, and all interactions related to galectin/LN1 complexes formation are colored in yellow.

doi:10.1371/journal.pone.0125946.g002



**Fig 3. Water-mediated interactions of hGal1 and hGal3 with the LN2 molecules.** (A and B) Stereoview of LN2 molecule bound in the carbohydrate-recognition site of hGal1 (PDB ID: 1W6P) and hGal3-CRD (1KJL), respectively. LN2 ligand is depicted as gray stick model. The water (blue sphere) is coordinated by the H-bonds (green dashes) from N2 atom of LN2 and amino acid residues of galectins.  $2F_o - F_c$  omit electron density (gray mesh) of the water molecules are highlighted and contoured at  $1\sigma$ . Unique salt bridge network of hGal1 and hGal3-CRD are indicated as yellow dashes. (C and D) Structural superposition of LN1 and LN2 complex structures from hGal1 (C) and hGal3-CRD (D). The C5-hydroxyl group of LN1 (yellow sticks) in hGal1 and hGal3-CRD complexes makes close contact with the coordinated water in LN2-hGal1 and LN2-hGal3-CRD complex with a distance of 2.2 and 2.1 Å, respectively. (E) Structure of hGal7 (shown in color green) in complex with LN2 (orange) is superimposed with LN2-hGal3-CRD complex structure (all in gray). Most LN2-contacting residues in hGal7 (such as Arg53<sup>hGal7</sup>, Trp69<sup>hGal7</sup>, Glu72<sup>hGal7</sup> and Arg74<sup>hGal7</sup>) are well superimposed with those of hGal3, except the Glu58<sup>hGal7</sup> residue. Location of Glu58<sup>hGal7</sup> is distinctive from Glu165<sup>hGal3</sup>/Asp54<sup>hGal1</sup> and distance between Glu58<sup>hGal7</sup> and N2 atom of LN2 is too far for them to coordinate a water molecule in between.

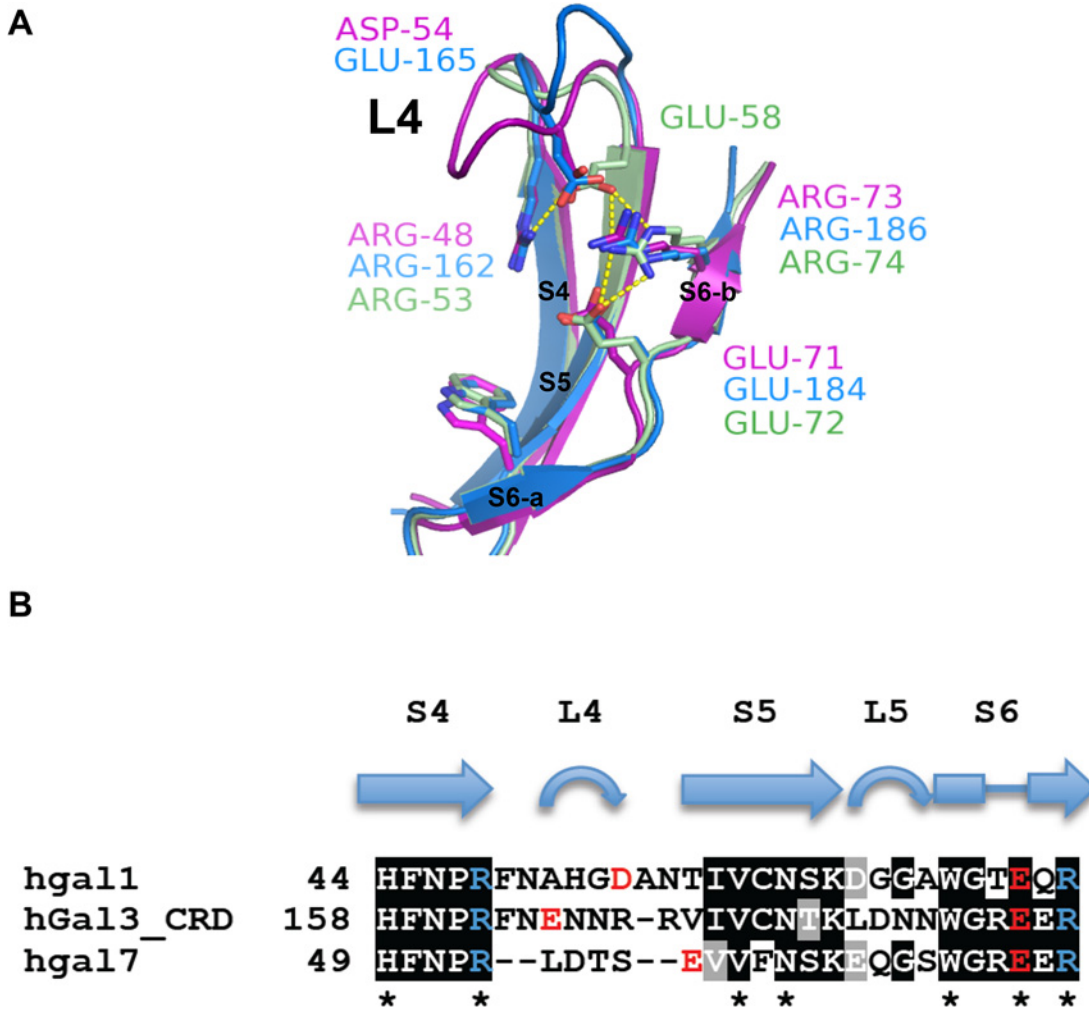
doi:10.1371/journal.pone.0125946.g003

LN2-interacting residues Arg53 and Glu72) to mediate the salt bridge network (Fig 2G), while both monodentate (Arg48<sup>hGal1</sup>-Asp54<sup>hGal1</sup> and Arg162<sup>hGal3</sup>-Glu165<sup>hGal3</sup>) and bidentate (Asp54<sup>hGal1</sup>-Arg73<sup>hGal1</sup> and Glu165<sup>hGal3</sup>-Arg186<sup>hGal3</sup>) interactions are observed in hGal1 and 3 (Fig 2E and 2F, respectively). Except for the residues involved in the unique salt-bridge network, especially those in hGal1 (Asp54<sup>hGal1</sup>-Glu71<sup>hGal1</sup>-Arg73<sup>hGal1</sup>) and hGal7 (Glu58<sup>hGal7</sup>-Glu72<sup>hGal7</sup>-Arg74<sup>hGal7</sup>) characteristic of larger RMSD (all atoms) (0.98 and 0.64 Å, respectively, see Fig 2E-2G), the majority of LN1/2-contact residues of hGal1, 3 and 7, such as His44<sup>hGal1</sup>/158<sup>hGal3</sup>/49<sup>hGal7</sup>, Asn46<sup>hGal1</sup>/160<sup>hGal3</sup>/51<sup>hGal7</sup>, Arg48<sup>hGal1</sup>/162<sup>hGal3</sup>/53<sup>hGal7</sup>, Val59<sup>hGal1</sup>/172<sup>hGal3</sup>/60<sup>hGal7</sup>, Asn61<sup>hGal1</sup>/174<sup>hGal3</sup>/62<sup>hGal7</sup> and Trp68<sup>hGal1</sup>/181<sup>hGal3</sup>/69<sup>hGal7</sup> are highly structural conserved with much smaller RMSD (all atoms) of 0.25, 0.13 and 0.29 Å, respectively (Fig 2A-2C).

Moreover, water-mediated interactions, identified to exist in the Asp54<sup>hGal1</sup>, Arg73<sup>hGal1</sup> and Glu165<sup>hGal3</sup>-situated salt-bridge networks (Fig 3A and 3B), were proposed to correlate with the LN2-binding specificity. A water was held by Asp54<sup>hGal1</sup>, Arg73<sup>hGal1</sup> and the N2 atom of GlcNAc in the hGal1/LN2 complex via formation of H-bonds (Fig 3A), whereas Glu165<sup>hGal3</sup> and the N2 of GlcNAc were observed to clamp a water molecule in the hGal3/LN2 complex (Fig 3B). The similar interactions were also found in the hGal1- and hGal3-lactose complexes (PDB IDs: 1W6O [14] and 2NN8 [21]) where the O2 atom of glucose was comparable to the N2 of GlcNAc in the previous two structures. However, there are no such water-mediated interactions in our hGal1-LN1 and hGal3-LN1 structures. As aforementioned, N-acetyl group in LN2 is structurally equivalent to C5-hydroxymethyl group in LN1 (Fig 2D). To accommodate an extra water molecule becomes impossible owing to the limited space for C5-hydroxymethyl group of LN1 and the residues participating in the salt-bridge network (Fig 3C and 3D). Meanwhile, neither of our hGal7-LN1 and the hGal7-LN2 complexes (PDB ID: 5GAL) [16] contains the aforementioned water-mediated interactions (Fig 3E). The main cause is the location of Glu58<sup>hGal7</sup> (Fig 3E) that is distinctive from that of Asp54<sup>hGal1</sup> and Glu165<sup>hGal3</sup>. Apparently Glu58<sup>hGal7</sup> was relatively remote from the bound sugar in hGal7, as compared to the analogous Asp54 in hGal1 and Glu165 in hGal3.

Superimposition of LN1-complexed hGal1, 3 and 7 indicated an important difference in the loop (denoted as L4) between S4 and S5 β-strands (Fig 4), and that Asp54<sup>hGal1</sup>, Glu165<sup>hGal3</sup> and Glu58<sup>hGal7</sup> are differently positioned in L4. L4 of hGal7 is shorter than the counterpart in hGal1 and hGal3 (Fig 4B). The additional amino acid residues in the L4 of hGal1 and 3 thus throng round the space in the vicinity of their carbohydrate-binding sites (Fig 4A). Unlike Asp54<sup>hGal1</sup> and Glu165<sup>hGal3</sup> both situated in the internal of L4, Glu58<sup>hGal7</sup> is resided in either the end of L4 or the beginning of S5 β-strand (Fig 4A and 4B), leading to a differently oriented salt bridge network from those of hGal1 and hGal3 (Fig 2E-2G). Such an arrangement makes it impossible for Glu58<sup>hGal7</sup> to coordinate with the N2 atom of LN2 for additional water-mediated interactions (Fig 3E). Taken together, the LN2-binding preference is possibly linked to the presence of water-mediated interactions, which is under the control of the properly positioned salt-bridge in L4.

Even Asp54<sup>hGal1</sup> and Glu165<sup>hGal3</sup> have a dissimilar location in L4 (Fig 4B), i.e. Asp54<sup>hGal1</sup> is resided at the end of the β-turn structure in L4, while Glu165<sup>hGal3</sup> appears as the first residue of the β-turn. Since Asp54<sup>hGal1</sup> and Glu165<sup>hGal3</sup> residues mediate ionic interactions to their adjacent Arg residues, apparently the location of Asp54<sup>hGal1</sup> and Glu165<sup>hGal3</sup> results in the distinct conformation of L4 in hGal1 and hGal3-CRD (Fig 4A). This explains the reason why hGal3, but not hGal1, binds with tumor-related TF antigen (Galβ1-3GalNAc) [40]. When the x-ray structure of hGal3-TF antigen is superimposed with the hGal3-LN1 structure, the galactose moiety is roughly overlapped ( $\phi = -90^\circ$  vs.  $-65^\circ$ , respectively), but the adjacent sugar residue adopts a very different orientation ( $\psi = 75^\circ$  vs.  $136^\circ$ ). The observation again demonstrates that,



**Fig 4. Structural comparisons among hGal1, 3-CRD and 7.** (A) S4-S6  $\beta$ -strands of hGal1 (pink), 3-CRD (blue) and 7 (green) are superimposed. The unique salt bridge network of hGal1 (R48-D54-E71-R73), 3-CRD (R162-E165-E184-R186) and 7 (R53-E58-E72-R74) are shown in stick models. (B) Structure-based sequence alignment of S4-S6  $\beta$ -strands of hGal1, 3-CRD and 7. Secondary structures were designated according to the resolved x-ray structures. The highly conserved LNs-interacting residues among hGal1, 3-CRD and 7 are indicated by asterisks. Residues involved in unique salt bridge network of hGal1, 3-CRD and 7 are colored in either red (Glu/Asp) or blue (Arg).

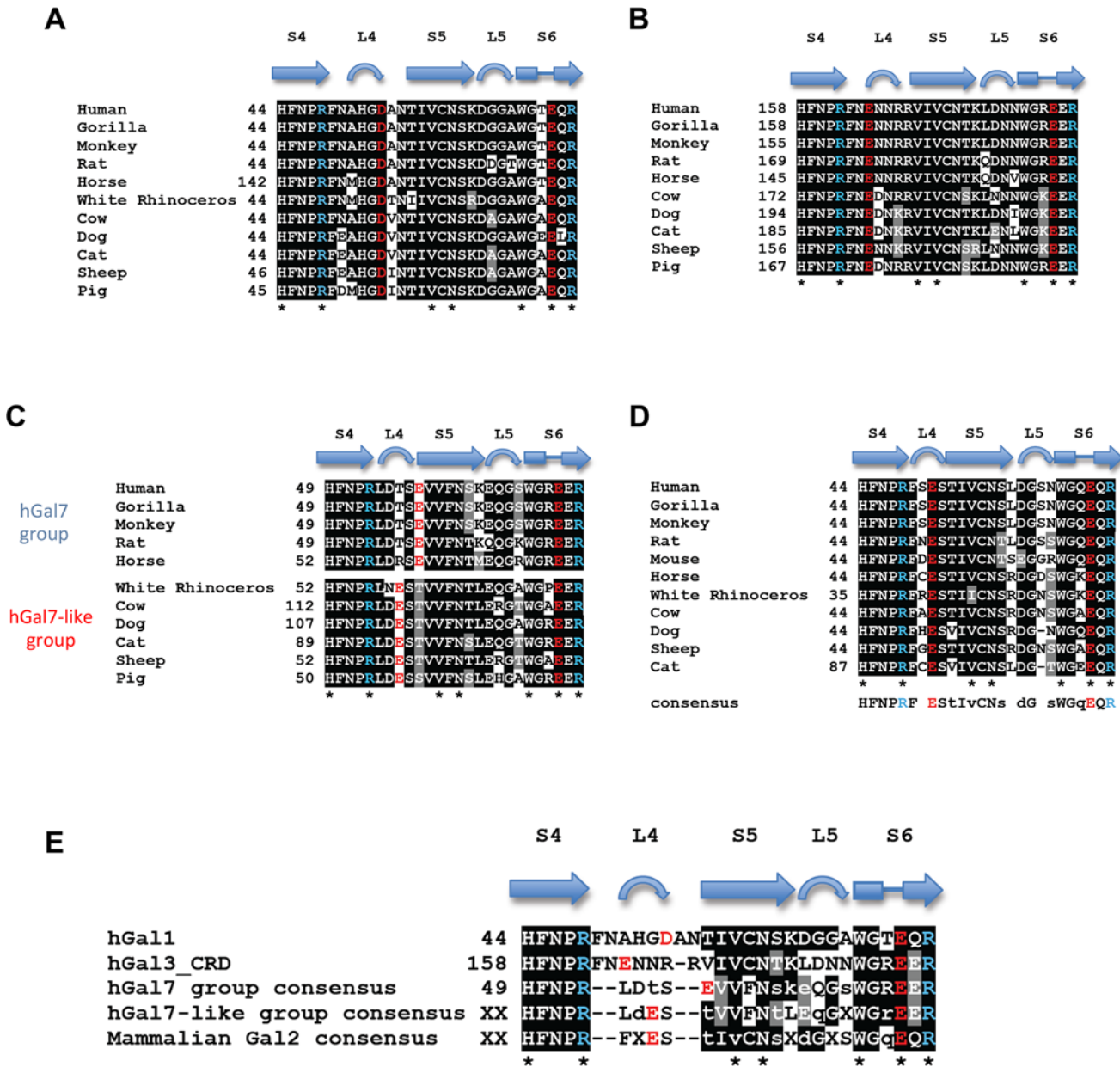
doi:10.1371/journal.pone.0125946.g004

in addition to the primary interactions with the galactose residue, the neighboring sugar has to rotate the glycosidic bond to interact well with several important residues, such as Arg162<sup>hGal3</sup> and Glu184<sup>hGal3</sup>. Although Glu165<sup>hGal3</sup> in L4 is not involved in the binding, this residue forms the salt-bridge network to connect with Arg162<sup>hGal3</sup> and Glu184<sup>hGal3</sup>.

Salt bridge motifs are known to play a key role in a number of functions, such as stabilization of protein folds, and arrangement of key residues or waters for the purpose of catalysis or molecular recognition [39, 41–43]. In this work the unique salt bridge motifs mediated by Asp54<sup>hGal1</sup>, Glu165<sup>hGal3</sup> and Glu58<sup>hGal7</sup> were identified to coordinate with the vicinal carbohydrate-binding sites to distinguish the bound sugar ligands, which accounts for the LN1/2-binding preferences of hGal1, 3 and 7, respectively.

As discussed previously, the sequence alignment of hGal1, 3 and 7 indicates that L4 is highly variable. But L4 of each galectin appears to be highly conserved among mammalian species (Fig 5), suggesting that L4 is likely to be evolved into different structures. Each of them is well





**Fig 5. Conservation of the L4 regions among mammalian galectins.** Structure-based sequence alignment of S4-S6  $\beta$ -strands of mammalian galectin-1s (A), galectin-3s (B), galectin-7s (C) and galectin-2s (D). According to the variation at corresponding position of Glu58<sup>hGal7</sup>, mammalian galectin-7s are further divided into two subgroups, the hGal7 group and hGal7-like group. Sequence comparison of L4 regions of human Galectin-1, 3 and mammalian Galectin-2, 7 (E). Invariant LNs-contacting residues of galectins are indicated by asterisks. Residues involved in unique salt-bridge network are colored according to negative (red) or positive (blue) charged properties. The sequences of mammalian galectins were selected from human (*Homo sapiens*: galectin-1 [NP\_002296.1], galectin-2 [NP\_006489.1], galectin-3 [BAA22164.1] and galectin-7 [NP\_002298.1]), Gorilla (*Gorilla gorilla gorilla*: galectin-1 [XP\_004063482.1], galectin-2 [XP\_004063485.1], galectin-3 [XP\_004055252.1] and galectin-7 [XP\_004060726.1]), Monkey (*Macaca mulatta*: galectin-1 [NP\_001162098.1], galectin-2 [XP\_001087063.1], galectin-3 [NP\_001253292.1] and galectin-7 [NP\_001083444.1]), Rat (*Rattus norvegicus*: galectin-1 [NP\_063969.1], galectin-2 [NP\_598283.1], galectin-3 [NP\_114020.1] and galectin-7 [NP\_072104.2]), Horse (*Equus caballus*: galectin-1 [XP\_001501082.2], galectin-2 [XP\_001499566.2], galectin-3 [XP\_005605252.1] and galectin-7 [XP\_001496714.2]), Rhinoceros (*Ceratotherium simum simum*: galectin-1 [XP\_004418181.1], galectin-2 [XP\_004418176.1] and galectin-7 [XP\_004441551.1]), Cow (*Bos taurus*: galectin-1 [NP\_786976.1], galectin-2 [NP\_001244020.1], galectin-3 [NP\_001095811.1] and galectin-7 [XP\_002695023.2]), Dog (*Canis lupus familiaris*: galectin-1 [ADR80617.1], galectin-2 [NP\_001271396.1], galectin-3 [NP\_001183972.1] and galectin-7 [NP\_001183972.1]), Cat (*Felis catus*: galectin-1 [XP\_003989294.1], galectin-2 [XP\_006934157.1], galectin-3 [XP\_003987704.1] and galectin-7 [XP\_003987704.1]), Sheep (*Ovis aries*: galectin-1 [AAT38511.1], galectin-2 [XP\_004007664.1], galectin-3 [XP\_004010713.1] and galectin-7 [XP\_004010713.1]), Pig (*Sus scrofa*: galectin-1 [NP\_001001867.1], galectin-3 [NP\_001090970.1] and galectin-7 [NP\_001136315.1]) and Mouse (*Mus Musculus*: galectin-2 [NP\_079898.2]).

doi:10.1371/journal.pone.0125946.g005

arranged by heavy H-bond connections between carbonyl and amide groups on their polypeptide backbones, supporting the previous hypothesis that different architectures and dynamics of these variable L4 regions might be functionally relevant for the carbohydrate-binding specificities and thus influence the biological properties of each galectin member [16, 40, 44–46]. Of particular note, the unique salt bridge networks of hGal1 (Arg48<sup>hGal1</sup>-Asp54<sup>hGal1</sup>-Glu71<sup>hGal1</sup>-Arg73<sup>hGal1</sup>) and hGal3 (Arg162<sup>hGal3</sup>-Glu165<sup>hGal3</sup>-Glu184<sup>hGal3</sup>-Arg186<sup>hGal3</sup>) are highly conserved (Fig 5A and 5B), while those in Gal7s are divided into two divergent groups due to the variations in the position of Glu58<sup>hGal7</sup> (Fig 5C). One group of mammalian Gal7s has the corresponding Glu in the end of L4 (hGal7 group), while the other group has this Glu shifted into the middle of L4 (hGal7-like group). Whether the hGal7-like group has a similar LN1-binding preference requires further investigations.

Moreover, it was reported that Gal1 and Gal2 are two most closely related members in the prototype galectin subfamily, and they share up to 43% amino acid sequence identity, in comparison with the 32% sequence identity between hGal1 and hGal7 [44, 46, 47]. In accordance with the analysis of Hirabayashi *et al.*, both rat galectin-2 and hGal7 displays prominent LN1-preferred binding activity [12]. Multiple sequence alignment of mammalian Gal2s revealed that they are highly conserved (Fig 5D) and characterized with a shorter L4 sequence that is reminiscent of hGal7-like group (Fig 5E). These analyses strongly suggested that the variable L4 region of galectins is highly relevant to the LN1- or LN2-preferred binding activity. To further demonstrate such relationship, alanine mutation was introduced to replace E165 or R186 of hGal3. Although the binding affinity of hGal3-E165A for LN1 and LN2 was reduced 2.5- and 8.5-fold, respectively (Table 1), the ratios of  $K_d^{LN1} / K_d^{LN2}$  were changed from 0.35 of wild-type hGal3 to 1.22 of hGal3-E165A (Table 4), suggesting that the original LN2-preferred binding of hGal3 was shifted to LN1-preference. Obviously the L4-directing Asp/Glu residue is indispensable for affecting LN1/2 binding preference. By mediating distinct salt-bridge network and water-mediated interaction, these charged residues at L4 serve as a key structural element to fine-tune the carbohydrate recognition. Nevertheless, the other site-directed mutant hGal3-R186A was found to lose its binding ability for both LN1 and LN2 although it was prepared with success and was no structurally deviated from the wild type hGal3. The lost binding was realized owing to the close relationship of Arg186<sup>hGal3</sup> with Glu184<sup>hGal3</sup> (a critical LN1/2-contact residue) [34, 48]. Furthermore, Bonzi *et al.* identified a conformational change of Arg73<sup>hGal1</sup> (corresponding to Arg186<sup>hGal3</sup> and Arg74<sup>hGal7</sup>) that was induced by the presence of a unique peptide, λ5-UR of pre-B cell receptor (pre-BCR) [49], leading to the modified

**Table 4. LN1-binding preference for hGal1, 3 and 7.**

Protein	LN1/LN2 ratio <sup>1</sup>	
	Current study	Previous study <sup>4</sup>
hGal1	0.44 <sup>2</sup> (0.53 <sup>3</sup> )	0.79
hGal3	0.35 <sup>2</sup> (0.23 <sup>3</sup> )	0.65 <sup>5</sup>
hGal7	1.52 <sup>2</sup> (2.97 <sup>3</sup> )	2.7
hGal3-E165A	1.22 <sup>2</sup>	

<sup>1</sup>To quantitatively evaluate LN1- or LN2-binding preference of hGal1, 3 and 7, LN1/LN2 ratios ( $K_d^{LN1} / K_d^{LN2}$ ) were calculated based on the  $K_d$  values presented in Table 1.

<sup>2</sup>The values were obtained based on the  $K_d^{BI}$  values determined by Biolayer Interferometry.

<sup>3</sup>The values were obtained based on the  $K_d^{FP}$  values determined by FP-based competition assays.

<sup>4</sup>It was reported by Hirabayashi, J. *et al.* [12].

<sup>5</sup>The CRD domain was used instead of full-length hGal3.

doi:10.1371/journal.pone.0125946.t004

carbohydrate-binding specificity of hGal1 (e.g. a 3-fold decrease for Lacto-N-neotetraose (LNnT) and a 10-fold increase for  $\alpha$ 3-SiaLacNAc). The binding change is crucial for pre-B cell maturation [49, 50], suggesting that the distinct salt-bridge network of galectin members are not only relevant to LN1- or LN2-preference, but also critical to determine the binding preference of other ligands.

## Conclusions

In summary, our structural studies of hGal1, hGal3-CRD and hGal7 explain how galectins exhibit the binding preference for Gal $\beta$ 1-3/4GlcNAc disaccharides. Since the galactose moiety affords primary interactions, the GlcNAc has to adopt different orientations for keeping comparable H-bonds with several Arg and Glu/Asp residues in a salt-bridge network. Because of the Glu/Asp residues in the variable loop L4, the length of L4 and the location of the Glu/Asp are found to influence the geometry of the salt bridge, resulting in the LN1/2-binding preference.

## Supporting Information

**S1 Fig. Octet Red biolayer interferometry-based equilibrium analysis for LN1 and LN2 binding to tip-bound hGal1, 3 and 7, respectively.** Steady state  $K_d$  values and statistical parameters of the fitting were listed by fitting curve from biolayer interferometry experiments at 27°C. Experimental procedures were detailed as those described in Materials and Methods. (TIF)

**S2 Fig. Fluorescence polarization analysis of (A) hGal1 (0–250  $\mu$ M), (B) hGal3 (0–230  $\mu$ M) and (C) hGal7 (0–360  $\mu$ M) with LN2-FITC (0.1  $\mu$ M) at 4°C.** Duplicate measurements are shown for all data points. The curves represent fitting of the data by nonlinear regression to the simplified formula for a one to one interaction by Prism 5.0 software (GraphPad, San Diego, CA). Their  $K_d$  values were extracted from the fitting and shown as indicated, respectively. (TIF)

## Acknowledgments

We thank Drs. Shang-Te Danny Hsu, Meng-Chiao Ho and Nei-Li Chan for discussion and/or critical reading of the manuscript and Drs. Meng-Ru Ho and Chris Jao of the Biophysics Core Facility for technical assistance. Portions of this research were carried out at beamlines BL13B1 and BL13C1 of the National Synchrotron Radiation Research Center (Taiwan) and beamline SP12B2 of the Spring8 (Japan).

## Author Contributions

Conceived and designed the experiments: TJH CHL. Performed the experiments: TJH HYL ZT SCW. Analyzed the data: TJH HYL SCW CHL. Contributed reagents/materials/analysis tools: TJH HYL ZT BSH CHL. Wrote the paper: TJH HYL CHL.

## References

1. Barondes SH, Castronovo V, Cooper DN, Cummings RD, Drickamer K, Feizi T, et al. Galectins: a family of animal beta-galactoside-binding lectins. *Cell*. 1994 Feb 25; 76(4):597–8. PMID: [8124704](#)
2. Leffler H, Carlsson S, Hedlund M, Qian Y, Poirier F. Introduction to galectins. *Glycoconjugate journal*. 2004; 19(7–9):433–40. PMID: [14758079](#)
3. Rabinovich GA, Toscano MA, Jackson SS, Vasta GR. Functions of cell surface galectin-glycoprotein lattices. *Curr Opin Struct Biol*. 2007 Oct; 17(5):513–20. PMID: [17950594](#)

4. Yang RY, Liu FT. Galectins in cell growth and apoptosis. *Cell Mol Life Sci*. 2003 Feb; 60(2):267–76. PMID: [12678492](#)
5. Liu FT, Rabinovich GA. Galectins as modulators of tumour progression. *Nat Rev Cancer*. 2005 Jan; 5(1):29–41. PMID: [15630413](#)
6. Stanley P, Cummings RD. Structures Common to Different Glycans. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, et al., editors. *Essentials of Glycobiology*. 2nd ed. Cold Spring Harbor (NY)2009.
7. Tateno H, Toyota M, Saito S, Onuma Y, Ito Y, Hiemori K, et al. Glycome diagnosis of human induced pluripotent stem cells using lectin microarray. *J Biol Chem*. 2011 Jun 10; 286(23):20345–53. doi: [10.1074/jbc.M111.231274](#) PMID: [21471226](#)
8. Hasehira K, Tateno H, Onuma Y, Ito Y, Asashima M, Hirabayashi J. Structural and quantitative evidence for dynamic glycome shift on production of induced pluripotent stem cells. *Mol Cell Proteomics*. 2012 Dec; 11(12):1913–23. doi: [10.1074/mcp.M112.020586](#) PMID: [23023295](#)
9. Urashima T, Fukuda K, Messer M. Evolution of milk oligosaccharides and lactose: a hypothesis. *Animal: an international journal of animal bioscience*. 2012 Mar; 6(3):369–74. doi: [10.1017/S1751731111001248](#) PMID: [22436215](#)
10. Sparrow CP, Leffler H, Barondes SH. Multiple soluble beta-galactoside-binding lectins from human lung. *J Biol Chem*. 1987 May 25; 262(15):7383–90. PMID: [2438278](#)
11. Mandal TK, Mukhopadhyay C. Binding free energy calculations of galectin-3-ligand interactions. *Protein Eng*. 2002 Dec; 15(12):979–86. PMID: [12601137](#)
12. Hirabayashi J, Hashidate T, Arata Y, Nishi N, Nakamura T, Hirashima M, et al. Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. *Biochim Biophys Acta*. 2002 Sep 19; 1572(2–3):232–54. PMID: [12223284](#)
13. Echeverria I, Amzel LM. Disaccharide binding to galectin-1: free energy calculations and molecular recognition mechanism. *Biophys J*. 2011 May 4; 100(9):2283–92. doi: [10.1016/j.bpj.2011.03.032](#) PMID: [21539798](#)
14. López-Lucendo MF, Solís D, André S, Hirabayashi J, Kasai KI, Kaltner H, et al. Growth-regulatory Human Galectin-1: Crystallographic Characterisation of the Structural Changes Induced by Single-site Mutations and their Impact on the Thermodynamics of Ligand Binding. *J Mol Biol*. 2004; 343(4):957–70. PMID: [15476813](#)
15. Sorme P, Arnoux P, Kahl-Knutsson B, Leffler H, Rini JM, Nilsson UJ. Structural and thermodynamic studies on cation- $\pi$  interactions in lectin-ligand complexes: high-affinity galectin-3 inhibitors through fine-tuning of an arginine-arene interaction. *J Am Chem Soc*. 2005 Feb 16; 127(6):1737–43. PMID: [15701008](#)
16. Leonidas DD, Vatzaki EH, Vorum H, Celis JE, Madsen P, Acharya KR. Structural basis for the recognition of carbohydrates by human galectin-7. *Biochemistry*. 1998 Oct 6; 37(40):13930–40. PMID: [9760227](#)
17. Sorme P, Kahl-Knutson B, Wellmar U, Nilsson UJ, Leffler H. Fluorescence polarization to study galectin-ligand interactions. *Methods Enzymol*. 2003; 362:504–12. PMID: [12968384](#)
18. Sorme P, Kahl-Knutsson B, Hullejt M, Nilsson UJ, Leffler H. Fluorescence polarization as an analytical tool to evaluate galectin-ligand interactions. *Anal Biochem*. 2004 Nov 1; 334(1):36–47. PMID: [15464951](#)
19. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. In: Charles W. Carter Jr., editor. *Methods Enzymol*: Academic Press; 1997. p. 307–26.
20. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica Section D, Biological crystallography*. 2010 Feb; 66(Pt 2):213–21. doi: [10.1107/S0907444909052925](#) PMID: [20124702](#)
21. Collins PM, Hidari KI, Blanchard H. Slow diffusion of lactose out of galectin-3 crystals monitored by X-ray crystallography: possible implications for ligand-exchange protocols. *Acta crystallographica Section D, Biological crystallography*. 2007 Mar; 63(Pt 3):415–9. PMID: [17327679](#)
22. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. *Acta crystallographica Section D, Biological crystallography*. 2011 Apr; 67(Pt 4):235–42. doi: [10.1107/S0907444910045749](#) PMID: [21460441](#)
23. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta crystallographica Section D, Biological crystallography*. 2010 Apr; 66(Pt 4):486–501. doi: [10.1107/S0907444910007493](#) PMID: [20383002](#)
24. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.3r1. 2010.

25. Concepcion J, Witte K, Wartchow C, Choo S, Yao D, Persson H, et al. Label-free detection of biomolecular interactions using BioLayer interferometry for kinetic characterization. *Combinatorial chemistry & high throughput screening*. 2009 Sep; 12(8):791–800.
26. Wartchow CA, Podlaski F, Li S, Rowan K, Zhang X, Mark D, et al. Biosensor-based small molecule fragment screening with biolayer interferometry. *J Comput Aided Mol Des*. 2011 Jul; 25(7):669–76. doi: [10.1007/s10822-011-9439-8](https://doi.org/10.1007/s10822-011-9439-8) PMID: [21660516](https://pubmed.ncbi.nlm.nih.gov/21660516/)
27. Saint-Lu N, Oortwijn BD, Pegon JN, Odouard S, Christophe OD, de Groot PG, et al. Identification of galectin-1 and galectin-3 as novel partners for von Willebrand factor. *Arterioscler Thromb Vasc Biol*. 2012 Apr; 32(4):894–901. doi: [10.1161/ATVBAHA.111.240309](https://doi.org/10.1161/ATVBAHA.111.240309) PMID: [22267483](https://pubmed.ncbi.nlm.nih.gov/22267483/)
28. Munoz FJ, Santos JI, Arda A, Andre S, Gabius HJ, Sinisterra JV, et al. Binding studies of adhesion/growth-regulatory galectins with glycoconjugates monitored by surface plasmon resonance and NMR spectroscopy. *Organic & biomolecular chemistry*. 2010 Jun 28; 8(13):2986–92.
29. Salomonsson E, Larumbe A, Tejler J, Tullberg E, Rydberg H, Sundin A, et al. Monovalent interactions of galectin-1. *Biochemistry*. 2010 Nov 9; 49(44):9518–32. doi: [10.1021/bi1009584](https://doi.org/10.1021/bi1009584) PMID: [20873803](https://pubmed.ncbi.nlm.nih.gov/20873803/)
30. Brewer CF. Thermodynamic binding studies of galectin-1, -3 and -7. *Glycoconjugate journal*. 2004; 19(7–9):459–65. PMID: [14758079](https://pubmed.ncbi.nlm.nih.gov/14758079/)
31. Tsai CM, Guan CH, Hsieh HW, Hsu TL, Tu Z, Wu KJ, et al. Galectin-1 and galectin-8 have redundant roles in promoting plasma cell formation. *J Immunol*. 2011 Aug 15; 187(4):1643–52. doi: [10.4049/jimmunol.1100297](https://doi.org/10.4049/jimmunol.1100297) PMID: [21753146](https://pubmed.ncbi.nlm.nih.gov/21753146/)
32. Houzelstein D, Reyes-Gomez E, Maurer M, Netter P, Higuera D. Expression patterns suggest that despite considerable functional redundancy, galectin-4 and -6 play distinct roles in normal and damaged mouse digestive tract. *J Histochem Cytochem*. 2013 May; 61(5):348–61. doi: [10.1369/0022155413478612](https://doi.org/10.1369/0022155413478612) PMID: [23360694](https://pubmed.ncbi.nlm.nih.gov/23360694/)
33. Dings RP, Miller MC, Nesmelova I, Astorgues-Xerri L, Kumar N, Serova M, et al. Antitumor agent calixarene 0118 targets human galectin-1 as an allosteric inhibitor of carbohydrate binding. *J Med Chem*. 2012 Jun 14; 55(11):5121–9. doi: [10.1021/jm300014q](https://doi.org/10.1021/jm300014q) PMID: [22575017](https://pubmed.ncbi.nlm.nih.gov/22575017/)
34. Hirabayashi J, Kasai K. Effect of amino acid substitution by sited-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa beta-galactoside-binding lectin. *J Biol Chem*. 1991 Dec 15; 266(35):23648–53. PMID: [1721052](https://pubmed.ncbi.nlm.nih.gov/1721052/)
35. Saraboji K, Hakansson M, Genheden S, Diehl C, Qvist J, Weininger U, et al. The carbohydrate-binding site in galectin-3 is preorganized to recognize a sugarlike framework of oxygens: ultra-high-resolution structures and water dynamics. *Biochemistry*. 2012 Jan 10; 51(1):296–306. doi: [10.1021/bi201459p](https://doi.org/10.1021/bi201459p) PMID: [22111949](https://pubmed.ncbi.nlm.nih.gov/22111949/)
36. Liao DI, Kapadia G, Ahmed H, Vasta GR, Herzberg O. Structure of S-lectin, a developmentally regulated vertebrate beta-galactoside-binding protein. *Proceedings of the National Academy of Sciences of the United States of America*. 1994 Feb 15; 91(4):1428–32. PMID: [8108426](https://pubmed.ncbi.nlm.nih.gov/8108426/)
37. Leffler H, Barondes SH. Specificity of binding of three soluble rat lung lectins to substituted and unsubstituted mammalian beta-galactosides. *J Biol Chem*. 1986 Aug 5; 261(22):10119–26. PMID: [3733704](https://pubmed.ncbi.nlm.nih.gov/3733704/)
38. Vidal P, Roldos V, Fernandez-Alonso Mdel C, Vauzeilles B, Bleriot Y, Canada FJ, et al. Conformational selection in glycomimetics: human galectin-1 only recognizes syn-Psi-type conformations of beta-1,3-linked lactose and its C-glycosyl derivative. *Chemistry*. 2013 Oct 18; 19(43):14581–90. doi: [10.1002/chem.201301244](https://doi.org/10.1002/chem.201301244) PMID: [24105715](https://pubmed.ncbi.nlm.nih.gov/24105715/)
39. Donald JE, Kulp DW, DeGrado WF. Salt bridges: geometrically specific, designable interactions. *Proteins*. 2011 Mar; 79(3):898–915. doi: [10.1002/prot.22927](https://doi.org/10.1002/prot.22927) PMID: [21287621](https://pubmed.ncbi.nlm.nih.gov/21287621/)
40. Bian CF, Zhang Y, Sun H, Li DF, Wang DC. Structural basis for distinct binding properties of the human galectins to Thomsen-Friedenreich antigen. *PloS one*. 2011; 6(9):e25007. doi: [10.1371/journal.pone.0025007](https://doi.org/10.1371/journal.pone.0025007) PMID: [21949831](https://pubmed.ncbi.nlm.nih.gov/21949831/)
41. Vetriani C, Maeder DL, Tolliday N, Yip KS, Stillman TJ, Britton KL, et al. Protein thermostability above 100 degreesC: a key role for ionic interactions. *Proceedings of the National Academy of Sciences of the United States of America*. 1998 Oct 13; 95(21):12300–5. PMID: [9770481](https://pubmed.ncbi.nlm.nih.gov/9770481/)
42. Bae E, Phillips GN Jr. Structures and analysis of highly homologous psychrophilic, mesophilic, and thermophilic adenylate kinases. *J Biol Chem*. 2004 Jul 2; 279(27):28202–8. PMID: [15100224](https://pubmed.ncbi.nlm.nih.gov/15100224/)
43. Bosshard HR, Marti DN, Jelesarov I. Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings. *Journal of molecular recognition: JMR*. 2004 Jan-Feb; 17(1):1–16. PMID: [14872533](https://pubmed.ncbi.nlm.nih.gov/14872533/)
44. Houzelstein D, Goncalves IR, Fadden AJ, Sidhu SS, Cooper DN, Drickamer K, et al. Phylogenetic analysis of the vertebrate galectin family. *Mol Biol Evol*. 2004 Jul; 21(7):1177–87. PMID: [14963092](https://pubmed.ncbi.nlm.nih.gov/14963092/)



45. Feng L, Sun H, Zhang Y, Li DF, Wang DC. Structural insights into the recognition mechanism between an antitumor galectin AAL and the Thomsen-Friedenreich antigen. *FASEB J*. 2010 Oct; 24(10):3861–8. doi: [10.1096/fj.10-159111](https://doi.org/10.1096/fj.10-159111) PMID: [20530247](https://pubmed.ncbi.nlm.nih.gov/20530247/)
46. Guardia CM, Gauto DF, Di Lella S, Rabinovich GA, Marti MA, Estrin DA. An integrated computational analysis of the structure, dynamics, and ligand binding interactions of the human galectin network. *Journal of chemical information and modeling*. 2011 Aug 22; 51(8):1918–30. doi: [10.1021/ci200180h](https://doi.org/10.1021/ci200180h) PMID: [21702482](https://pubmed.ncbi.nlm.nih.gov/21702482/)
47. Lohr M, Lensch M, Andre S, Kaltner H, Siebert HC, Smetana K Jr, et al. Murine homodimeric adhesion/growth-regulatory galectins-1, -2 and -7: comparative profiling of gene/ promoter sequences by database mining, of expression by RT-PCR/immunohistochemistry and of contact sites for carbohydrate ligands by computational chemistry. *Folia biologica*. 2007; 53(4):109–28. PMID: [17706016](https://pubmed.ncbi.nlm.nih.gov/17706016/)
48. Abbott WM, Feizi T. Evidence that the 14 kDa soluble beta-galactoside-binding lectin in man is encoded by a single gene. *Biochem J*. 1989 Apr 1; 259(1):291–4. PMID: [2719646](https://pubmed.ncbi.nlm.nih.gov/2719646/)
49. Elantak L, Espeli M, Boned A, Bornet O, Bonzi J, Gauthier L, et al. Structural basis for galectin-1-dependent pre-B cell receptor (pre-BCR) activation. *J Biol Chem*. 2012 Dec 28; 287(53):44703–13. doi: [10.1074/jbc.M112.395152](https://doi.org/10.1074/jbc.M112.395152) PMID: [23124203](https://pubmed.ncbi.nlm.nih.gov/23124203/)
50. Bonzi J, Bornet O, Betzi S, Kasper BT, Mahal LK, Mancini SJ, et al. Pre-B cell receptor binding to galectin-1 modifies galectin-1/carbohydrate affinity to modulate specific galectin-1/glycan lattice interactions. *Nature communications*. 2015; 6:6194. doi: [10.1038/ncomms7194](https://doi.org/10.1038/ncomms7194) PMID: [25708191](https://pubmed.ncbi.nlm.nih.gov/25708191/)