CHEMMEDCHEM

CHEMISTRY ENABLING DRUG DISCOVERY

Accepted Article

Title: Structure-based Design of a Monosaccharide Ligand Targeting Galectin-8

Authors: Mohammad Bohari, Xing Yu, Chandan Kishor, Brijesh Patel, Rob-Marc Go, Hadieh Alsadat Eslampanah Seyedi, Yaron Vinik, Irwin Darren Grice, Yehiel Zick, and Helen Blanchard

To be cited as: *ChemMedChem* 10.1002/cmdc.201800224

Link to VoR: http://dx.doi.org/10.1002/cmdc.201800224

A Journal of

ChemPubSoc Europe

WILEY-VCH

www.chemmedchem.org

FULL PAPER

WILEY-VCH

Structure-based Design of a Monosaccharide Ligand Targeting Galectin-8

Mohammad H. Bohari [a], Xing Yu [a], Chandan Kishor [a], Brijesh Patel [a], Rob Marc Go [a], Hadieh A. Eslampanah Seyedi [a], Yaron Vinik [b], I. Darren Grice [a,c], Yehiel Zick [b], and Helen Blanchard [a]*

[a] Dr. M. H. Bohari (ORCID: 0000-0001-9375-5290), Dr. X. Yu, Dr. C. Kishor, B. Patel, R. M. Go, H. A. E. Seyedi, Prof. H. Blanchard (ORCID: 0000-0003-3372-5027)
Institute for Glycomics
Griffith University
Parkland's Drive, Gold Coast Campus 4222, Australia
E-mail: h.blanchard@griffith.edu.au

[b] Dr. Y. Vinik, Prof. Y. Zick
Department of Molecular Cell Biology,
Weizmann Institute of Science
Herzl St 234, Rehovot, Israel.

[c] Dr. I. D. Grice (ORCID: 0000-0001-9218-457X)
School of Medical Science
Griffith University
Parkland's Drive, Gold Coast Campus, 4222, Australia.

Supporting information for this article is given via a link at the end of the document.

Abstract: Galectin-8 is a β-galactoside-recognising protein that has a role in the regulation of bone remodelling and is an emerging new target for tackling diseases with associated bone-loss. We have designed and synthesised methyl 3-*O*-[1-carboxyethyl]-β-D-galactopyranoside (compound **6**) as a ligand to target the *N*-terminal domain of galectin-8 (galectin-8*N*). Our design involved molecular dynamics (MD) simulations that predicted **6** to mimic the interactions made by the galactose ring as well as the carboxylic acid group of 3′-*O*-sialylated lactose (3′-SiaLac), with galectin-8*N*. Isothermal titration calorimetry (ITC) determined that the binding affinity of galectin-8*N* for **6** was 32.8 μM, while no significant affinity was detected for the *C*-terminal domain of galectin-8 (galectin-8*C*). The crystal structure of the galectin-8*N*-**6** complex validated the predicted binding conformation and revealed the exact protein-ligand interactions that involve galectin evolutionarily conserved amino acids and also those unique to galectin-8*N* for recognition. Overall, we have initiated and demonstrated a rational ligand design campaign to develop a monosaccharide-based scaffold as a binder of galectin-8.

Introduction

Galectin-8 is a β-galactoside recognising protein that contains two carbohydrate recognition domains (CRD) in tandem, linked by a variable length amino acid linker [1]. It has widespread tissue distribution in both normal and tumour cells. Within the cell, galectin-8 occurs in the nucleus, the cytoplasm, and also it is secreted into the extracellular space [2]. Apart from being involved in cell-to-cell and cell-to-surrounding communication, an increasingly broader functional spectrum of galectin-8 is apparent [3]. Galectin-8 plays an important role in inflammatory disorders through the regulation of T-cell homoeostasis [4] and is critically involved in capillary tube formation and angiogenesis [5]. Antibacterial activity, mediated through induction of selective autophagy, highlights an additional cellular mechanism to combat infection recruiting galectin-8 [6]. It is of interest that galectin-8 has shown *in vivo* regulation of bone remodelling via enhancing the expression of bone resorbing factors that are attributed to increased bone turnover culminating in reduced bone mass [7]. Inhibition of galectin-8 may thus offer a potential new therapeutic approach in managing diseases with bone-loss. Available structural information of galectin-8*N* bound to different biologically relevant oligosaccharides provide insight into various biological processes and highlights key interactions imparting affinity and specificity to a ligand [8]. Typically, the CRD has a β-sandwich "jelly-roll" topology formed from two β-sheets. The concave surface of the roll that constitutes the carbohydrate-binding site, comprises six β-strands, S1-S6 (Figure 1a). The galactose recognition site consists of evolutionarily conserved amino acids on strands S4-S6. Galectin-8*N* preferentially

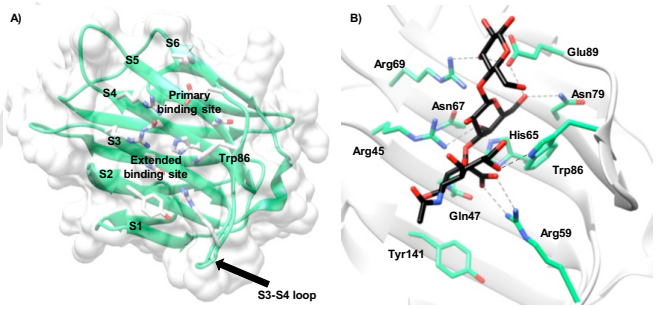


Figure 1. Overview of the galectin-8*N* carbohydrate recognition domain. a) The CRD showing the carbohydrate binding face of the "jelly-roll". The primary and the extended binding site are indicated. b) Binding conformation and hydrogen bond and salt-bridge interactions (in grey dashed lines) made by 3′-SiaLac (in sticks; carbon in black, oxygen in red) upon binding to galectin-8*N* residues (in sticks; carbon in green, oxygen in red and nitrogen in blue) PDB ID: 3AP7 [8a].

recognises anionic oligosaccharides such as 3′-sulfated lactose and 3′-*O*-sialylated lactose (3′-SiaLac) [9]. This preferential binding arises from the presence of unique structural features in the galectin-8*N* CRD such as a long S3-S4 loop, bearing an arginine residue (Arg59), and a glutamine (Gln47) on strand S3 [8a]. The crystal structure of 3′-SiaLac bound to galectin-8*N* shows engagement of the carboxylic acid group of sialic acid in salt bridge interactions with Arg59 and hydrogen bonding to

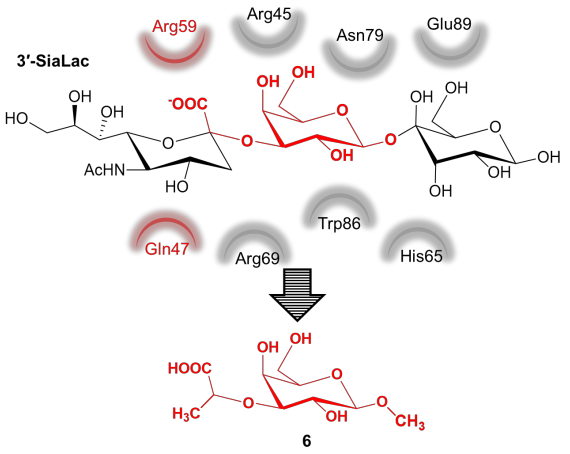


Figure 2. The ligand design concept. Depicted are the chemical structures of 3′-SiaLac, the designed compound **6,** and the surrounding amino acid residues in the galectin-8*N* binding site. The interaction with Arg59 and Gln47 (labelled in red) are the unique residues that **6** is designed to exploit.

Gln47, and the pyranose ring of the galactose formed conserved interactions that are observed for the counterpart of lactose (Figure 1b) [8a]. These interactions between the anionic group of the glycans with Arg59 and Gln47 were attributed as the origin of the high affinity towards galectin-8*N* [8a].

Given the inherent nature of galectins to recognise β-galactoside-containing glycans, and galectin-8*N's* preferential recognition of anionic sugars, then a monosaccharide-based ligand such as galactose bearing a negatively charged carboxylic acid group at the 3′-position akin to 3′-SiaLac seemed promising as a ligand scaffold. Incorporation of a carboxylic acid group in ligands to engage unique binding site residues of galectin-8*N* was also supported by our *in silico* virtual screening of a library of compounds with a diverse set of functional groups where 25% of the top hits identified had a carboxylic acid group (*unpublished*). The mammalian galectins have wide ranging biological functions but exhibit a shared primary binding site and hence exploiting unique residue/s in close proximity of this site is vital in achieving galectin specificity. We report the first synthetic ligand (methyl 3-*O*-[1-carboxyethyl] -β-D-galactopyranoside, see Figure 2) (**6**) developed through a structure-based rational design approach targeting galectin-8 (specifically the *N*-terminal CRD). We demonstrate its ability to bind with affinity at the micromolar range to a monosaccharide-based scaffold, as well as its preference towards the galectin-8*N* domain.

Results and Discussion

Ligand design hypothesis though MD simulations

MD simulations were performed to analyse the relative strength and propensity of binding interactions observed in the crystal structures, as well as for detection of possible alternate interactions assisting in conceptualising further our ligand design hypothesis. Lactose as the native galectin ligand, was also included (galectin-8*N*-lactose complex PDB ID: 5T7S [8b]) in the MD analysis along with the galectin-8*N*-3′-SiaLac crystal structure (PDB ID: 3AP7 [8a]). Hydrogen bond occupancy analysis was performed to examine the longevity of interactions observed in the crystal structure. The hydrogen bonds observed in galectin-8*N*-lactose and galectin-8*N*-3′-SiaLac crystal structures were found with occupancy of 60-100 % (total occupancy for the residue) during simulations (Figure 3). An identical interaction profile for the common lactose moiety within both ligands was observed for the conserved residues His65, Glu89, Arg69, and Asn79 (Figure 3). The occupancy for Arg45 hydrogen bonds was higher in 3′-SiaLac compared to that in the lactose complex. The carboxylate group of 3′-SiaLac showed 100 % occupancy of salt bridge interactions with Arg59 and over 60 % occupancy of hydrogen bond with Gln47 and Trp86. Our simulation of galectin-8*N*-3′-SiaLac crystal complex revealed that 60 % of the time the Gln47 side chain was re-orientated by 180° interchanging the $N^{\delta1}$ and $O^{\delta1}$ compared to the deposited PDB coordinates (PDB ID: 3AP7[8a]) and thus enabled hydrogen bonding to occur with the carboxylic acid group. Critically, amino acids Arg59 and Gln47 are unique to galectin-8*N* and are identified as supposed ligand specificity hotspots [8]. Taking advantage of this existing structural information and findings from our *in silico* virtual screening (not repoted here) a monosaccharide-based ligand of galectin-8*N* was conceived to exploit interactions with both unique and conserved residues of the galectin-8*N* binding site.

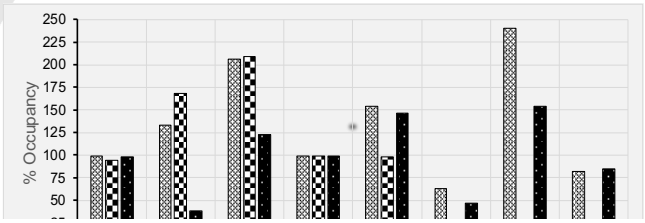
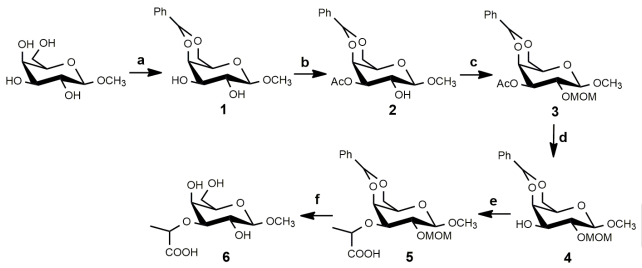


Figure 3. Interaction analysis from 100 ns MD simulations. Percentage of frames with specific hydrogen bonds between the galectin-8*N* binding site residues and the ligand (3′-SiaLac, lactose (Lac) and **6**) atoms. For clarity, the hydrogen bond occupancy with respect to individual residues is represented.

The stable interactions made by the galactose ring and the carboxylic acid group in the galectin-8*N*-3′-SiaLac crystal complex were incorporated in the design of **6** (Figure 2). The main advantage for **6** being the anticipated specific recognition by galectin-8*N,* and associated increase in binding affinity and specificity, by the carboxylic acid group on the 3′-position of the galactose. Furthermore, **6** can be synthesised with relative ease and has increased stability over 3′-SiaLac. Lactose (molecule's atomic coordinates) from the galectin-8*N*-lactose complex (PDB ID: 5T7S [8b]) was modified to obtain the initial placement of **6** in the galectin-8*N* binding site (Supplementary Figure S1). This *in silico* galectin-8*N*-**6** complex was subjected to 100 ns simulation

to investigate the feasibility of the interactions observed in the initial model and the suitability of **6** in the binding site. Compound **6** stayed in the primary binding site for the duration of the simulation, retaining the typical CH-π interactions with the evolutionarily conserved Trp86, as observed for the natural galectin ligands. Hydrogen bond occupancy analysis revealed an almost identical interaction profile of **6** to that observed for the corresponding part of the 3′-SiaLac in the galectin-8*N*-3′-SiaLac simulation (Figure 3). The hydrogen bond occupancy, particularly for the unique residues Arg59 and Gln47 and the conserved residue Trp86, noted for **6** is identical to that observed for 3′-SiaLac. The MD interaction analysis suggests that the galactose ring of **6** would occupy the primary binding site of galectin-8*N* and that the carboxylic acid group at the 3′-position would engage in interactions with the unique Arg59 and Gln47. Based on these predictions, we synthesised **6**, and its binding affinity, binding mode and interactions to galectin-8*N* were validated using ITC and X-ray crystallography.

Chemical synthesis of the designed ligand



Scheme 1. **Synthesis scheme for 6.** Reagents and condition: a) CSA, benzaldehyde dimethyl acetal, ACN, 60ºC, 2.5 h (yield 88%); b) Ag_2O (freshly prepared), AcCl, KI, DCM, RT, 16 h (yield 70%); c) DIPEA, MOMBr, DCM, reflux, 16 h (yield 80%); d) Na metal, MeOH, RT, 1.5 h (yield 90%); e) NaH, 2-chloropropionic acid, anhydrous 1,4-dioxane, 50 ºC, 36 h (yield 60%); f) concentrated HCl, MeOH (yield 75 %).

Synthesis of **6** was initiated with β-D-methyl galactoside as the starting material (Scheme 1). The C4-OH and C6-OH group of the β-D-methyl galactoside was 4,6-benzylidene protected (**1**) utilising benzaldehyde dimethyl acetal with a catalytic amount of camphor sulfonic acid (CSA)[10] to afford **1**. Following this, **1** underwent selective 3*O*-acetylation using silver oxide (Ag_2O), acetyl chloride (AcCl) and a catalytic amount of potassium iodide (KI)[11] to give **2**. The acetylated galactoside **2** was then 2*O*-methoxymethyl ether (MOM) protected using an adapted procedure [12], employing diisopropyl ethyl amine (DIPEA) in dichloromethane (DCM) to give **3**. After quantitative deacetylation using sodium metal in methanol to give **4**, the 2-chloropropionic acid side chain was coupled using sodium hydride (NaH) in 1,4-dioxane to yield **5** [13]. A final deprotection at positions 2- and 4,6-positions was performed using concentrated HCl in methanol, to yield a racemic mixture of compound **6** [14]. The structure of **6** was confirmed with the ^{13}C NMR signature chemical shifts for two methyl groups being observed, one consistent with the anomeric at 55.84 ppm and the other for the propionic acid side chain at 17.63 ppm. Additionally, corresponding ^{1}H NMR peaks were observed and a molecular ion from MS confirmed the integrity of the ligand.

Validation of interactions by X-ray crystal structure determination

We subsequently performed crystallographic analysis to investigate the binding mode and interactions of **6** to galectin-8*N*. The galectin-8*N*-**6** complex was obtained by soaking the ligand into the *apo* galectin-8*N* crystals, and the structure then determined at 2.1 Å resolution. The electron density maps reveal unambiguous placement of **6** in the galectin-8*N* binding site (Figure 4, Supplementary Figure S2). A clear protrusion in the difference electron density map pointing towards the conserved Arg69 confirmed the positioning of the O4′ hydroxyl of the galactose. Further, the planar topology of the electron density adjacent to the unique Arg59 is consistent with the placement of the carboxylic acid group and thereby confirming the overall placement of **6** (Figure 4). The electron density for the methyl group of the propionic acid side chain that points toward solvent is weak. However, it is sufficient to give the direction of methyl group, correlating with an *R*-configuration at the carbon chiral centre. The positive difference electron density that appears as an extension from the anomeric methoxy group, is concluded to be a water molecule.

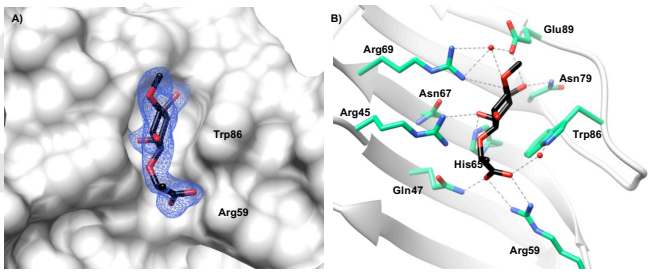


Figure 4. Galectin-8*N*-6 complex. A) Electron density (blue mesh) 2|Fo|-|Fc| α_c contoured at 1σ, for **6** (in sticks; carbon in black, oxygen in red) in complex with galectin-8*N*. B) Hydrogen bonding and salt-bridge interactions (in grey dashed line) made between **6** and galectin-8*N* (in sticks; carbon in green; oxygen in red; nitrogen in blue).

Overall, the binding mode observed for the galactose portion of **6** is identical to that noted for the corresponding part of the galectin-8*N*-lactose complex [8b]. The O4′ of the galactose engages in hydrogen bonding with His65, Asn67, Arg45, Arg69 whereas the O6′ hydrogen bonds with Asn79, and Glu89, as also noted from our simulations (Figure 4). Importantly, from the original design concept, the carboxylic acid was found in a geometrically favoured position to make ionic interactions with the unique Arg59 and hydrogen bonding interactions with Gln47. Furthermore, the placement of the anionic group and the galactose ring is identical to the equivalent part of 3′-SiaLac complexed with galectin-8*N* (Supplementary Figure S3). The carboxylic acid group displaced a water molecule that was observed in the vicinity of the conserved Trp86 and Arg59 in the galectin-8*N*-lactose complex (PDB ID: 5T7S [8b]), suggesting overall stronger binding. With the galectin-8*N*-**6** crystal structure, we validated our design concept and the predicted binding conformation for **6**.

Table 1. Crystallographic data merging and refinement statistics for galectin-8*N*-**6** complex structure.

Data	Galectin-8*N*-**6**
Indexing	
Crystal system	Orthorhombic
Space group	P212121
Unit cell	a=45.72, b=50.32, c=80.87
Merging and Scaling	
Resolution (Å)	42.73 – 2.10
Total observations	54110 (4147) [a]
Unique observations	11126 (873)
Multiplicity	4.9 (4.8)
Completeness (%)	98.0 (95.0)
I/σ	5.9 (1.8)
Rmerge (%)	14.9 (78.2)
Refinement	
Resolution	42.59-2.10
Rfactor (%)	22.2
Rfree (%)	25.1
Number of atoms	
Protein	1196
Ligand	18
Water	103
Root mean square deviation	
Bond length (Å)	0.007
Bond angle (°)	1.29
Ramachandran plot statistics	
Favoured (%)	97.2
Allowed (%)	2.8
Average B-factor (Å^2)	
Protein	26.3
Ligand	36.3
Water	33.0
PDB ID	5VWG

[a] The value in parenthesis is for the highest-resolution shell.

Binding specificity through MD simulations

Assessing the preference of **6** towards the CRDs of galectin-8, we have employed MD simulations to predict and compare its binding mode and interactions with galectin-8*N* and galectin-8*C*. The primary binding site of the two CRDs of the galectin-8 is mostly conserved. However, amino acid differences in the extended binding site would play a critical role in recognising **6**. We used the crystallographic conformation of **6** from the galectin-8*N*-**6** complex (PDB ID: 5VWG) and generated the *in silico* galectin-8*C*-**6** complex model through superimposition of the two CRDs. The *in silico* galectin-8*C*-**6** complex model reveals possible hydrogen bonding interactions of **6** with His271 (His65 in galectin-8*N*), Arg275 (Arg69), Glu294 (Glu89), and Asn284 (Asn79) as observed in the galectin-8*N*-**6** crystal structure (Figure 5). However, the carboxylic acid side chain of **6** lacked interactions with the galectin-8*C* binding site due to the absence of Arg59 on S3-S4 loop, the presence of Ser255 in place of Arg45, and Asn257 in place of Gln47 (Figure 5). The favourability and stability of ligand binding were analysed from100 ns MD simulations, while binding free energies were estimated using the MMPBSA method. Compound **6** occupied the primary binding site in galectin-8*N* during the length of simulation, with no significant fluctuations in the ligand placement detected (Figure 5). However, **6** showed significant fluctuations in case of the galectin-8*C* binding site, possibly due to lack of the long S3-S4 loop bearing unique residue like Arg59 (Figure 5). The average estimated ligand binding free energy for the galectin-8*C*-**6** complex was therefore observed to be only half (-25.5 kcal/mol) that estimated for the galectin-8*N*-**6** complex (-60.6 kcal/mol). Overall, our simulation results and binding free energy analysis suggest favourability of the designed compound **6** towards galectin-8*N* compared to galectin-8*C*.

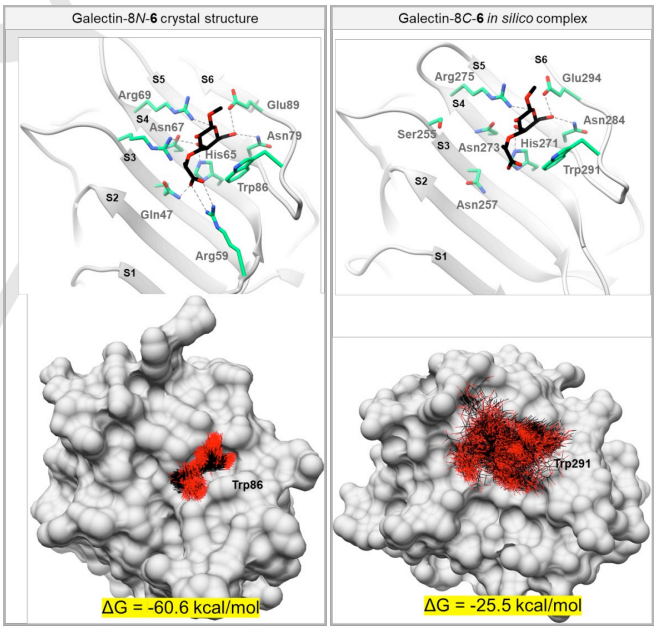


Figure 5: Binding mode comparison of 6 towards the galectin-8*N* (crystal structure [PDB ID: 5VWG]; on the left) and the galectin-8*C* (*in silico* complex; on the right). On top, key binding site residues are represented in sticks (carbon in green for protein atoms and carbon in black for ligand atoms) and hydrogen bonds and salt-bridge in grey dashed lines. Bottom figures show overlayed ligand (in the wire; carbon in black and oxygen in red) coordinates extracted from the 100 ns simulation bound to galectin-8*N* (bottom right) and galectin-8*C* (bottom left). The MMPBSA estimated binding free energy (in kcal/mol) is highlighted in yellow.

Binding affinity determination and specificity by ITC

Binding affinities (K_d) of compound **6**, lactose and 3′-SiaLac towards galectin-8*N* and galectin-8*C* were determined using ITC. In the experiment, heat changes occurring from the ligand's

titration into a protein solution and the ligand's titration into a buffer (without the protein) were recorded. Compound **6** exhibited 1:1 stoichiometric binding to galectin-8*N* with a binding affinity of 32.8 ± 1.8 μM, whilst that determined for 3′-SiaLac was 2.3 ± 0.2 μM (Figure 6 and Supplementary Figure S4). The binding affinity of the monosaccharide **6** was slightly stronger than the measured affinity for the disaccharide lactose (47.4 ± 1.1 μM), indicating the overall efficiency of the designed ligand. Interestingly, **6** did not show significant binding towards galectin-8*C*. This is presumed to be mainly due to the absence of Arg59 on the S3-S4 loop, supporting the observed difference in the estimated binding free energies from MD simulations. The

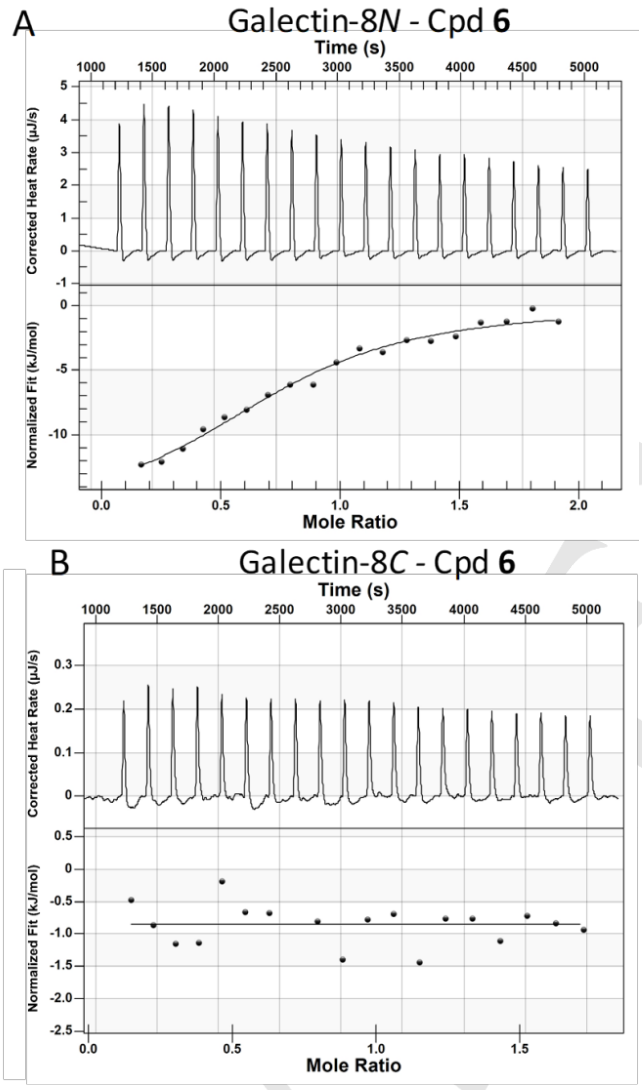


Figure 6. **Isothermal calorimetric analysis**. Binding isotherm for titration of 1mM of **6**. A) with 200 mM galectin-8*N* and B) 200 μM galectin-8*C* in Tris buffer at pH 8 containing 100 mM NaCl and 4 mM BME. Lactose and 3′-SiaLac ITC graphs are presented in supplementary data (Figure S4).

affinity of galectin-8*C* for lactose was 331.1 μM, whilst no significant binding was detected for 3′-SiaLac, as also previously reported [9]. These findings suggest the binding preference of **6** towards galectin-8*N* over galectin-8*C*. The results with compound **6** thus support our design hypothesis of exploring the unique residue Arg59 found in galectin-8*N* for gaining affinity along with specificity. Overall, our binding affinity data is in good agreement with the previously reported affinities for lactose and 3′-SiaLac towards galectin-8*N* and galectin-8*C* [9] (Table 2). The discrepancy in the absolute K_d values previously reported for lactose (1.7-3.1 μM) and 3′-SiaLac (53 nM) using the fluorescence anisotropy assay could potentially be due to the difference in the principle of detection and variation in experimental parameters such as buffer, pH and temperature [15].

Table 2. Binding affinity and estimated thermodynamic parameters from ITC experiments.

	Ligand	Kd (μM)	ΔH (kJ/mol)	-TΔS (kJ/mol)	ΔG (kJ/mol)	n
Galectin-8*N*	Lactose	47.4 ±1.1	-6.4 ±1.3	-18.3	-24.6	0.92 ±0.04
	3′-SiaLac	2.3 ±0.2	-52.7 ±0.8	20.5	-32.1	0.81 ±0.01
	Cpd **6**	32.8 ±2.1	-16.3 ±1.3	-9.3	-25.6	0.77 ±0.04
Galectin-8C	Lactose	331.1	-0.6 ± 0.6	-19.2	-19.8	1.28 ±0.28
	3′-SiaLac	>1000	-	-		-
	Cpd **6**	No binding	-	-		-

Conclusions

Overall, we have employed a structure-based approach to rationally design a monosaccharide-scaffold ligand to target galectin-8*N*. The preference of anionic saccharides towards galectin-8*N* [9, 15] and the results of our *in silico* virtual screening suggested the favourability of anionic groups in the galectin-8*N* binding site. Considering previously obtained *in silico* data, the most frequent hydrogen bonding interactions exhibited during MD simulations, the synthetic feasibility and relative stability over 3′-SiaLac, **6** was designed. The ligand design hypothesis was taken experimentally validated by synthesising **6** and evaluating its binding to galectin-8*N* through X-ray crystallography and ITC. The X-ray structure of the galectin-8*N*-**6** complex validated our predicted conformation, where the ligand explored both the evolutionarily conserved and the unique amino acids of galectin-8*N* for interaction. The interaction profile observed for **6** was identical to the corresponding part of native ligand lactose, and the template molecule 3-SiaLac when bound to galectin-8*N*. The binding specificity of **6** was addressed by determining its binding affinity using MD simulations and ITC towards the two CRDs of galectin-8. The ITC determined binding data suggests preferential binding of compound **6** to galectin-8*N* (32.8 μM) over galectin-8*C* (no binding detected), and its binding affinity was slightly better than the affinity of lactose (47.4 μM). This favourability was also indicated from our MD simulations and was most likely the result of our targeted exploitation of unique residues for interactions by **6**. The promising binding affinity of our monosaccharide-based **6** over the disaccharide

lactose is indicative of the overall improved ligand efficiency. Furthermore, this prompt for exploration of other functional groups on the galactose core towards identifying novel compounds as galectin-8*N* ligands and potential inhibitors. Our rational ligand-design campaign has led to successful identification of a monosaccharide-based scaffold that binds to galectin-8*N* and forms a proof-of-concept study. This scaffold provides a basis for further optimisation aimed at enhancing binding affinity and specificity towards galectin-8 with the aimed application as a therapeutic in managing diseases that have associated bone-loss.

Experimental Section

Molecular dynamics simulations

The initial coordinates of the designed compound 6 bound to galectin-8*N* were obtained by modifying the lactose from the galectin-8*N*-lactose complex (5T7S [8b]) using the BIOVIA Discovery Studio Visualiser [16]. *In silico* galectin-8*C*-6 complex was generated by removing the bound peptide and superimposing the galectin-8*C* crystal structure (4GXL [17]) on to the galectin-8*N*-6 crystal structure using the MatchMaker utility of UCSF Chimera [18]. All MD simulations were performed using the GROMACS 4.5.6 version [19] with AMBER99SB-ILDN force field [20], as employed previously [8b, 21]. Long-range electrostatics were handled using Particle Mesh Ewald method [22]. Ligand topology and parameters were generated by applying AM1-BCC charges and Generalised Amber Force Field [23] using a python script Acpype [24] that uses Antechamber module of AMBER [25]. The protein-ligand complexes were initially energy minimised using the steepest descent method, followed by position restrained minimisation and finally the 100 ns production run. Hydrogen bond analysis was performed using the g_hbond utility available with the GROMACS package, and the occupancy analysis was performed using the python script readHBmap, written by R. O. S. Soares. Visualisation of MD trajectories was carried out in VMD [26]. The ligand binding free energy was estimated using molecular mechanics energies with the Poisson-Boltzmann and surface area continuum solvation method (MMPBSA.py) [27] implemented in Amber package [28]. A set of 100 frames periodically extracted from the trajectory file at an interval of 1 ns were subjected to MMPBSA analysis to obtain the ligand binding free energies.

Chemical Synthesis

General procedure: Thin Layer Chromatography (TLC) on pre-coated aluminium-backed silica plates (60 F_{254}; Merck) were used to assess reaction progression by visualisation after charring in 4% sulphuric acid in ethanol. Reaction products were purified using flash chromatography silica gel 60. ^{1}H and ^{13}C NMR spectra were recorded at 298 K using an Avance Bruker Biospin spectrometer (400 and 100 MHz, respectively; Bruker Biospin) spectrometer. Two-dimensional COSY (^{1}H to ^{1}H correlation), HSQC (^{1}H to ^{13}C correlation) and HMBC (^{1}H to ^{13}C long range correlation) NMR experiments were used to assist in assigning relevant peaks for ^{1}H and ^{13}C NMR spectra. Electrospray ionisation low-resolution mass spectrometry was performed using a Bruker Daltronics Esquire 3000 Ion-Trap MS.

Methyl 4,6-O-benzilidene-β-D-galactopyranoside (1) [10] : Methyl-β-D-galactopyranoside (1 g, 5.1 mmol) was dried *in vacuo* overnight before being dissolved in anhydrous acetonitrile (CH_3CN) (10 mL) under argon. A catalytic amount of camphor sulfonic acid (CSA) (52 mg, 0.22 mmol) was added followed by dropwise addition of benzaldehyde dimethyl acetal (1.5 mL, 14.8 mmol) added dropwise. The reaction was heated to 60 °C for about 3 hours. The reaction was quenched with triethylamine (Et_3N), then purified via flash chromatography using 40:1 dichloromethane (DCM):methanol (MeOH) to yield 1 (88%). The product was consistent with the reported title compound [10]. ^{1}H NMR ($CDCl_3$, 400 MHz): δ = 3.52 (1H, s), 3.59 (3H, s, OCH_3), 4.05 (1H, dd, *J*=7.6, 11.6 Hz), 4.13 (1H, dd, *J*=1.6, 12.4 Hz), 4.35 (1H, d, *J*=7.6 Hz), 4.41 (1H, dd, *J*=1.6, 12.4 Hz), 4.46 (1H, dd, *J*=0.8, 3.6 Hz), 4.85 (1H, dd, *J*=3.68, 11.6 Hz), 5.50 (1H, s, $PhCHO_2$), 7.35 (3H, m, ArH), 7.49 (2H, m, ArH).

Methyl 3-O-acetyl 4,6-O-benzilidene-β-D-galactopyranoside (2) [11] : Compound 1 was dissolved in anhydrous DCM, cooled to -20 °C using an ice-salt mixture. Freshly prepared silver oxide (Ag_2O) [29] was added and left for 30 minutes with stirring, followed by slow addition of acetyl chloride (AcCl) and KI [11]. The reaction was left stirring overnight at room temperature. Ag_2O was removed via filtration, and the solvent was removed *in vacuo*. The product was purified by flash chromatography (1:1 hexane:ethyl acetate (EtOAc)) to give 2 (70% yield). ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 2.11 (3H, s, OAc), 2.52 (1H, brs), 3.48 (1H, d, *J*=1.5 Hz), 3.56 (3H, s, OCH_3), 3.93 (1H, dd, *J*= 1.6, 8.0), 4.01 (1H, m, *J*=2.0, 12.4 Hz), 4.21 (1H, d, *J*=8.0 Hz, H1), 4.26 (1H, dd, *J*=1.6, 12.4 Hz), 4.32 (1H, dd, *J*=0.8, 3.6 Hz), 4.82 (1H, dd, *J*=3.6, 10.2 Hz), 5.48 (1H, s, OCHPh), 7.32-7.37 (3H, m, ArH), 7.47-7.50 (2H, m, ArH).

Methyl 2-O-methoxymethyl-3-O-acetyl-4,6-O-benzilidene-β-D-galactopyranoside (3): The methoxy methyl (MOM) ether protection procedure was adapted from literature [12]. Compound 2 was dissolved in DCM under argon at room temperature, diisopropyl ethylamine (DIPEA) was added at 0°C followed by drop-wise addition of bromomethyl methyl ether and refluxed overnight. The reaction was diluted with DCM and washed with water and brine solution, then purified using flash column chromatography (2:1 hexane:EtOAc) to give 3 (80% yield). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 2.10 (3H, s, OAc), 3.38 (3H, s, MOM), 3.47 (1H, d, *J*=1.1 Hz), 3.55 (3H, s, OCH_3), 3.95 (1H, q, *J*=2.4, 10.2 Hz), 4.04 (1H, dd, *J*=1.8, 12.4 Hz), 4.31 (1H, dd, *J*=1.5 Hz), 4.33 (1H, d, *J*=2.2 Hz), 4.35 (1H, dd, *J*=0.7, 3.7 Hz), 4.68 (1H, d, *J*=6.4 Hz, H1), 4.85 (2H, m, MOM), 5.48 (1H, s, CHPh), 7.35 (3H, m, ArH), 7.50 (2H, m, ArH). ^{13}C NMR ($CDCl_3$, 100 MHz): δ = 21.1, 55.7, 56.9, 66.1, 69.0, 72.7, 73.5, 97.2, 101.1, 104.1, 126.4, 128.1, 129.0, 137.7, 170.8. MS (ESI): *m/z* calculated for $C_{18}H_{24}NaO_8$ $[M+Na]^+$ 391.2, found 391.2.

Methyl 2-O-methoxymethyl-4,6-O-benzilidene-β-D-galactopyranoside (4): Compound 3 was dissolved in MeOH and cooled to 0 °C before addition of sodium metal previously suspended in hexane. The reaction was left at room temperature for 1.5 hours, then carefully acidified to pH 5 using 1 M HCl. Salts were removed by water washing, and the product was extracted with EtOAc, then solvent removed *in vacuo* to give 4 (90% yield). ^{1}H NMR (CD_3OD, 400 MHz): δ 3.35 (3H, s, MOM), 3.47 (1H, s), 3.48 (3H, s, OCH_3), 3.56 (1H, dd, *J*= 7.6, 9.6 Hz), 3.62 (1H, dd, *J*= 3.6, 9.6 Hz), 4.06 (1H, d, *J*=1.6, 12.4 Hz), 4.13 (2H, m, *J*=1.6 Hz), 4.24 (1H, d, *J*=7.6 Hz), 4.66 (1H, d, *J*=6.4 Hz), 4.76 (2H, s, MOM), 5.53 (1H, s, CHPh), 7.33 (3H, m), 7.33 (3H, m), 7.49 (2H, m). ^{13}C NMR (CD_3OD, 100 MHz): δ = 54.7, 55.8, 66.5, 68.7, 71.8, 75.6, 76.3, 96.7, 101.1, 104.6, 126.0, 127.0, 128.5, 138.3. MS (ESI): *m/z* calculated for $C_{16}H_{22}NaO_7$ $[M+Na]^+$ 349.1, found 349.1.

Methyl 3-O-[1-carboxyethyl]-β-D-galactopyranoside (6): The propionic acid side chain was installed onto 4 using previously reported conditions [13]. Compound 4 was dissolved in anhydrous 1,4-dioxane under argon and cooled to 0°C before addition of NaH. 2-Chloropropionic acid was slowly added at 0 °C, then the reaction was stirred at 50 °C for 36 hours to yield the racemic mixture of novel ligand 5 (60% yield). ^{1}H NMR ($CDCl_3$, 400 MHz): δ = 1.64 (3H, d, *J*=6.8 Hz, C*H*$_3$CH), 3.36 (3H, s, OCH_3), 3.39 (1H, m), 3.48 (3H, s, OCH_3), 3.61 (1H, dd, *J*=3.2, 9.6 Hz), 3.77 (1H, dd, *J*=8.0, 10.0 Hz), 4.16 (2H, m, *J*=1.2, 12.4 Hz, H6), 4.38 (3H, m, *J*=8.0 Hz), 4.51 (1H, d, *J*=3.2 Hz), 4.78 (1H, d, *J*=6.4 Hz), 5.62 (1H, s, C*H*Ph), 7.25-7.34 (3H, m), 7.40-7.49 (2H, m). ^{13}C NMR (CD_3OD, 100 MHz): δ = 18.1, 19.2, 56.5, 56.0, 56.9, 57.0, 66.3, 65.9, 69.0, 69.1,72.5, 72.8, 73.6, 74.0, 75.2, 75.3, 79.5, 81.3, 97.5, 98.2, 101.1, 101.7, 103.7, 103.8, 126.3, 126.6, 128.2, 128.4, 129.1, 129.5, 136.8, 137.5, 173.6,

174.0. MS (ESI): *m/z* calculated for $C_{19}H_{26}NaO_9$ $[M+Na]^+$ 421.2, found 421.3.

The deprotection of 5 was carried out in MeOH and concentrated HCl at room temperature. The enantiomeric mixture of novel ligand 6 was purified using reversed-phase chromatography (C_{18}, 4:1 water:MeOH). ^{1}H NMR (CD_3OD, 400 MHz): δ = 1.23 (3H, d, *J*=8.0 Hz), 3.33 (1H, dd, *J*=3.2, 9.6 Hz), 3.41 (3H, s, OCH_3), 3.40-3.66 (4H, m), 3.92 (1H, m), 3.98 (1H, q, *J*=8.0 Hz, C*H*CH_3), 4.18 (1H, dd, *J*=6.0, 8.0 Hz). ^{13}C NMR (CD_3OD, 100 MHz): δ = 17.6, 55.8, 61.0, 67.2, 70.7, 74.9, 75.2, 82.1, 104.5. MS (ESI): *m/z* calculated for $C_{10}H_{18}KO_8$ $[M+K]^+$ 305.4, found 306.1.

Protein expression and purification

The galectin-8*N* protein was expressed in an untagged form as described previously [8b]. Briefly, the bacterial culture was induced with IPTG for 4 hours at room temperature and purified using affinity chromatography on a lactosyl-Sepharose column at 4 °C. The purity of the expressed protein was assessed using SDS-PAGE and was used directly for binding assays and crystallisation.

Galectin-8*C* was expressed with 10x His-tag by using pET-19b vector in the *E. coli* BL21(DE3) strain. The *E. coli.* cells were grown in Luria-Bertani (LB) medium supplemented with 100 µg/ml Ampicillin to reach an OD_{600} of 0.8. The culture was induced with 0.25mM IPTG and incubated at 16°C for 18 hrs to reach the optimum yield. To extract the cytoplasmic soluble protein, cells were lysed using sonication and centrifuged to harvest the supernatant. Galectin-8*C* was purified using affinity chromatography on Ni-NTA column equilibrated with PBS, pH 7.9, containing 10mM imidazole. Protein was then eluted using 370mM imidazole and dialyzed against Tris buffer with 2.5 mM β-mercaptoethanol (BME). The purified protein was concentrated using Amicon Ultra centrifugal filter units (Millipore, Billerica, MA, USA) with a 3kDa molecular weight cut-off (MWCO). The purity of the expressed protein and obtained protein fractions were assessed by SDS-PAGE. Protein concentration was determined using UV spectroscopy at 280 nm.

Isothermal titration calorimetry

Quantitation of the binding affinity was performed by measuring the dissociation constant using Nano_ITC Instrument (MicroCal,TA). Galectin-8*N* and galectin-8*C* were prepared in Tris buffer (20 mM Tris, pH 8.0, 100 mM NaCl and 4 mM BME). Samples were degassed prior to use. Titrations were performed in a TA NanoAnalyze calorimeter using 20 injections applied at an interval of 150 seconds at 298 K. Each injection dispensed 2.5 µl of ligand (Lactose, 3′-SiaLac and 6) into the sample cell containing 300 µl of either Galectin-8*N* or Galectin-8*C*, at 200 rpm. Ligand titration and blank data were collected at room temperature, and binding isotherms were fitted using an independent model in Nano_Analyze™ v3.7 software.

X-ray data collection and atomic structure determination

The galectin-8*N*-6 complex structure was obtained by soaking 6 into the *apo* galectin-8*N* crystals, as performed previously [8b]. The *apo* galectin-8*N* crystals were generated in phosphate buffer saline (10 mM sodium phosphate, 137 mM sodium chloride, 2.7 mM potassium chloride, 1.8 mM potassium phosphate; PBS). Compound 6 was dissolved in PBS at a concentration of 10 mM and soaked into the *apo* crystals for 5 mins. X-ray diffraction data were collected at room temperature. The crystals were mounted on the goniometer using the MicroRT capillary system from MiTigen. A Rigaku MicroMax™-007 HF rotating anode generator coupled with VariMax optics and shutter-less PILATUS 200 K detector was used to perform the experiments. HKL-3000R[30] was used to control the instrument, and HKL-3000R and iMosflm [31] were used for data processing including indexing, integration and scaling. PHASER [32] implemented in CCP4 [33] was used for molecular replacement with the *apo* galectin-8*N* (PDB ID: 3AP5 [8a]) structure used as the search model. The chemical information file for 6 was generated using the PRODRG server [34]. Refinement was carried out using REFMAC5 [35], model building, and visualisation done in COOT [36]. Final model validation performed using MolProbity [37]. Molecular graphics and electron density illustrations for figures were performed using the UCSF Chimera package [18, 38].

Acknowledgements

H.B gratefully acknowledges the financial support from the Cancer Council Queensland (ID1080845).

Accession code: Protein Data Bank: Atomic coordinates and structure factors have been deposited with PDB accession code 5VWG for the galectin-8*N*-methyl 3-*O*-[1-carboxyethyl]-β-D-galactopyranoside (**6**) complex.

Keywords: Structure-based ligand design • carbohydrate • galectin-8*N* terminal domain • galectin-8*C* terminal domain • X-ray crystal structure

References:

[1] a) Z. Z. Su, J. Lin, R. Shen, P. E. Fisher, N. I. Goldstein, P. B. Fisher, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*(14), 7252-7257; b) N. Bidon, F. Brichory, S. Hanash, P. Bourguet, L. Dazord, J. P. Le Pennec, *Gene* **2001**, *274*(1-2), 253-262.

[2] a) Y. R. Hadari, K. Paz, R. Dekel, T. Mestrovic, D. Accili, Y. Zick, *J. Biol. Chem.* **1995**, *270*(7), 3447-3453; b) R. V. Gopalkrishnan, T. Roberts, S. Tuli, D. Kang, K. A. Christiansen, P. B. Fisher, *Oncogene* **2000**, *19*(38), 4405-4416; c) N. Nagy, Y. Bronckart, I. Camby, H. Legendre, H. Lahm, H. Kaltner, Y. Hadari, P. Van Ham, P. Yeaton, J. Pector, Y. Zick, I. Salmon, A. Danguy, R. Kiss, H. Gabius, *Gut* **2002**, *50*(3), 392-401.

[3] a) N. Bidon-Wagner, J. P. Le Pennec, *Glycoconj. J.* **2004**, *19*(7-9), 557-563; b) Y. Zick, M. Eisenstein, R. A. Goren, Y. R. Hadari, Y. Levy, D. Ronen, *Glycoconj. J.* **2004**, *19*(7-9), 517-526.

[4] a) L. Eshkar Sebban, D. Ronen, D. Levartovsky, O. Elkayam, D. Caspi, S. Aamar, H. Amital, A. Rubinow, I. Golan, D. Naor, Y. Zick, I. Golan, *J. Immunol. (Baltimore, Md.: 1950)* **2007**, *179*(2), 1225-1235; b) J. F. Sampson, A. Suryawanshi, W. S. Chen, G. A. Rabinovich, N. Panjwani, *Immunol. Cell Biol.* **2016** 94(2), 213-219; c) A. Norambuena, C. Metz, L. Vicuna, A. Silva, E. Pardo, C. Oyanadel, L. Massardo, A. Gonzalez, A. Soza, *J. Biol. Chem.* **2009**, *284*(19), 12670-12679; d) M. V. Tribulatti, V. Cattaneo, U. Hellman, J. Mucci, O. Campetella, *J. Leukoc. Biol.* **2009**, *86*(2), 371-380.

[5] V. M. Delgado, L. G. Nugnes, L. L. Colombo, M. F. Troncoso, M. M. Fernandez, E. L. Malchiodi, I. Frahm, D. O. Croci, D. Compagno, G. A. Rabinovich, C. Wolfenstein-Todel, M. T. Elola, *FASEB J.* **2011**, *25*(1), 242-254.

[6] a) S. R. Stowell, C. M. Arthur, M. Dias-Baruffi, L. C. Rodrigues, J. P. Gourdine, J. Heimburg-Molinaro, T. Ju, R. J. Molinaro, C. Rivera-Marrero, B. Xia, D. F. Smith, R. D. Cummings, *Nat. Med.* **2010**, *16*(3), 295-301; b) T. L. Thurston, M. P. Wandel, N. von Muhlinen, A. Foeglein, F. Randow, *Nature* **2012**, *482*(7385), 414-418.

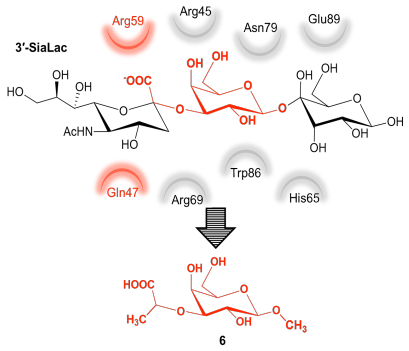
[7] Y. Vinik, H. Shatz-Azoulay, A. Vivanti, N. Hever, Y. Levy, R. Karmona, V. Brumfeld, S. Baraghithy, M. Attar-Lamdar, S. Boura-Halfon, I. Bab, Y. Zick, *eLife* **2015**, *4*, e05914.

[8] a) H. Ideo, T. Matsuzaka, T. Nonaka, A. Seko, K. Yamashita, *J. Biol. Chem.* **2011**, *286*(13), 11346-11355; b) M. H. Bohari, X. Yu, Y. Zick, H. Blanchard, *Sci. Rep.* **2016**, *6*, 39556.

[9] H. Ideo, A. Seko, I. Ishizuka, K. Yamashita, *Glycobiology* **2003**, *13*(10), 713-723.

[10] C.-H. Wong, F. Moris-Varas, S.-C. Hung, T. G. Marron, C.-C. Lin, K. W. Gong, G. Weitz-Schmidt, J. Am. Chem. Soc **1997**, *119*(35), 8152-8158.
[11] H. Wang, J. She, L.-H. Zhang, X.-S. Ye, *J. Org. Chem.* **2004**, *69*(17), 5774-5777.
[12] K.-i. Sato, K. Sakai, K. Tsushima, S. Akai, *Tet. Lett.* **2007**, *48*(21), 3745-3748.
[13] L. K. Lars Andersson, *Carbohydr. Res.* **1998**, *313*, 157-164.
[14] N. Hama, T. Aoki, S. Miwa, M. Yamazaki, T. Sato, N. Chida, *Org. Lett.* **2011**, *13*(4), 616-619.
[15] S. Carlsson, C. T. Oberg, M. C. Carlsson, A. Sundin, U. J. Nilsson, D. Smith, R. D. Cummings, J. Almkvist, A. Karlsson, H. Leffler, *Glycobiology* **2007**, *17*(6), 663-676.
[16] Dassault Systèmes BIOVIA, *San Diego: Dassault Systèmes, 2016.*
[17] S. Li, M. P. Wandel, F. Li, Z. Liu, C. He, J. Wu, Y. Shi, F. Randow, *Sci. Signal.* **2013**, *6*(261), ra9.
[18] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*(13), 1605-1612.
[19] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*.
[20] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw, *Proteins* **2010**, *78*(8), 1950-1958.
[21] P. M. Collins, K. Bum-Erdene, X. Yu, H. Blanchard, *J. Mol. Biol.* **2014**, *426*(7), 1439-1451.
[22] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys* **1995**, *103*(19), 8577-8593.
[23] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*.
[24] A. W. Sousa da Silva, W. F. Vranken, *BMC Res. Notes.* **2012**, *5*(1), 1-8.
[25] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. Model.* **2006**, *25*.
[26] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14*.
[27] a) P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, T. E. Cheatham, 3rd, *Acc. Chem. Res.* **2000**, *33*(12), 889-897; b) B. R. Miller, 3rd, T. D. McGee, Jr., J. M. Swails, N. Homeyer, H. Gohlke, A. E. Roitberg, *J. Chem. Theory Comput.* **2012**, *8*(9), 3314-3321.
[28] T. A. D. D.A. Case, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman, *University of California, San Francisco* **2012**.
[29] D. Y. Curtin, R. J. Harder, *J. Am. Chem. Soc* **1960**, *82*(9), 2357-2368.
[30] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307-326.
[31] T. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. Leslie, *Acta Crystallogr.* D. **2011**, *67*(Pt 4), 271-281.
[32] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*(Pt 4), 658-674.
[33] Collaborative Computational Project, Number 4. *Acta Crystallogr.* D. **1994**, *50*(Pt 5), 760-763.
[34] A. W. Schuttelkopf, D. M. van Aalten, *Acta Crystallogr.* D. **2004**, *60*(Pt 8), 1355-1363.
[35] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr.* D. **2011**, *67*(Pt 4), 355-367.
[36] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr.* D. **2010**, *66*(Pt 4), 486-501.
[37] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr.* D. **2010**, *66*(Pt 1), 12-21.
[38] T. D. Goddard, C. C. Huang, T. E. Ferrin, *J. Struct. Biol.* **2007**, *157*(1), 281-287.

Entry for the Table of Contents



Exploiting both evolutionarily conserved and unique binding-site amino acids in ligand design: Targeting human galectin-8, we have undertaken *in silico* structure-based design and chemical synthesis to generate a novel efficient ligand, and validated its binding by ITC and X-ray crystallography.