Exploring the *In situ* Pairing of Human Galectins and α-Dystroglycan *O*-Mannosylated Core M1 Glycopeptides

（コアM1型糖鎖を有するα-ジストログリカン糖ペプチドとガレクチンの*In-situ*ペア探索研究）

Doctoral Dissertation

2023 March

Lareno L. Villones Jr.

Laboratory of Advanced Chemical Biology

Graduate School of Life Science, Hokkaido University

Table of Contents

Abstract

Dystroglycan (DG), which constitutes a part of the dystrophin–glycoprotein complex, connects the extracellular matrix to the cytoskeleton. DG glycans are presented by the extracellular α-DG, serving as a contact point beyond the well-studied interaction between matriglycan and laminin G-like domains, providing muscular and neural cell stability. However, it remains unknown as to whether core M1 (GlcNAcβ1-2Man) structures can serve as ligands among the various *O*-Mannosylated (*O*-Man) glycans. On the other hand, galectin (Gal) is a family of carbohydrate-binding proteins (CBPs) that bind specifically to β-galactose-containing glycoconjugates modulating wide-range of (patho)physiological processes, such as cell growth/adhesion/differentiation, regulation of immune response, inflammatory function, and tumor development and progression. Therefore, based on the presence of *N*-acetyllactosamine (LacNAc) in this type of glycan following core extension, the binding interactions with adhesion/growth-regulatory galectins were explored.

To elucidate this process, the interaction between the galectin (Gal)-1, -3, -4 and -9 with the core M1-based glycopeptide library of the α-DG fragment ^{372}TRGAIIQTPTLGPIQPTRV^{390} were profiled, using glycan microarray. The binding of the galectins was revealed irrespective of the type of modular architecture (Gal-1>>Gal-4≈Gal-9, but very weak interaction with Gal-3), adding galectins to the list of possible binding partners of α -DG peptide and LacNAc presenting core M1 glycoconjugates by *cis*-binding via peptide- and carbohydrate-protein interactions, respectively. The binding of galectins was abrogated by α2,3-sialylation of the LacNAc units. This molecular event was further verified by nuclear magnetic resonance studies, wherein the LacNAc-terminated α-DG glycopeptide was found to simultaneously interact with both the S- and F-faces of Gal-1, thereby inducing oligomerization. Furthermore, the *trans*-bridging capabilities of Gal-1 with α-DG core M1 structures and laminins (-111, -121, -211, and -221, but little -511) were observed, which proposed the possible mechanism by which Gal-1 prevents muscular dystrophies; however, this proposal warrants further investigation.

The multifunctionality of galectins regulating a broad range of fundamental cellular processes via *cis*- and *trans*-activities has achieved wide attention in exploring beyond the importance of natural specificity/selectivity to the glycoconjugate receptors, i.e., its modular architecture to present the carbohydrate recognition domain (CRD) variously. Combining Gal-1, -3, -4, and -9 variant test panels and synthetic α-DG *O*-Man core M1 glycopeptide microarray library, a detailed comparative analysis is possible in delineating design-functionality relationships within this lectin family towards its affinity to the prepared glycoconjugates of α-DG. Enhancement of affinity towards the prepared ligands was observed in linker-connected di- and tetramer Gal-1 variants, while converting this to a Gal-3-like protein decreased binding. Presenting Gal-3 as a prototype markedly increased the susceptibility to the test compounds. While inserting a peptide linker between Gal-3 CRDs to form a tandem-repeat type diminished the binding. The galectin-4 variants have shown that the natural linker is detrimental to its interaction with α-dystroglycan glycoconjugates. On the other hand, Gal-9 variants revealed the importance of the C-terminal CRD on the binding affinity. Compared to Gal-1 wild-type, Gal-1 variants demonstrated higher *trans*-bridging capabilities between LacNAc- and sialyl-LacNAc-terminated *O*-Man core M1 α-dystroglycan glycopeptides and laminins (-111, -121, -211, -221, and -511) *in situ* (Fig. 2). This suggests possible higher translational applications of these galectin variants in the treatment of some forms of α-dystroglycanopathy.

Overall, our experimental setup revealed that *O*-Man core M1 glycopeptides of α-DG could serve as ligands for galectins in situ via cis-binding. In addition, the prototype Gal-1 can trans-bridge O-Man core M1 glycopeptides of α-DG and laminins in microarray. We also demonstrated that the alteration of the galectin structures can give additional insights into the preferential modular architecture and binding behavior of this lectin family towards specific ligands. Furthermore, rational protein engineering is indeed a useful tool in redesigning lectins with possibly higher therapeutic potentials than their wild-type counterpart. Here, Gal-1 and its variants has proof-of-principle character in *trans*-bridge *O*-Man core M1 glycopeptides of α-DG and laminins *in situ*.

Abbreviations

DG	Dystroglycan
α-DG	α-Dystroglycan
Gal	Galectin
CBP	Carbohydrate binding proteins
O-Man	*O*-Mannosylation
Core M1	(GlcNAcβ1-2Man)
LacNAc	*N*-acetyllactosamine
Sialyl-LacNAc	6-sialyl-*N*-acetyllactosamine
DTT	Dithiothreitol
Methyl-Lac	Methy-lactose
NMR	Nuclear Magnetic Resonance
HSQC	Heteronuclear Single Quantum Coherence
Lam	Laminin
FCMD	Fukuyama Congenital Muscular Dystrophy

Chapter 1.

General introduction

1.1. The sugar code

Carbohydrates are commonly known as a major source of energy (glycogen in animals), as structural components (chitin in insects and cellulose in plants), or as UV protective materials (ribose and deoxyribose in genes); that lack other significant biological activities (1). However, the latest findings in glycobiology have proved that sugar chains (glycans) attached to lipid and protein microdomains play an essential role as biological signals for cell-to-cell interaction and quality control of proteins that maintain tissue structure, porosity, and integrity (2). Thus, recently, carbohydrates are considered the third alphabet of life, next to nucleic acids and amino acids (first and second alphabet of life, respectively) (Fig. 1). During the recognition processes, these biopolymers are viewed as "messages", "read" by their corresponding receptors triggering a wide range of molecular events (3).

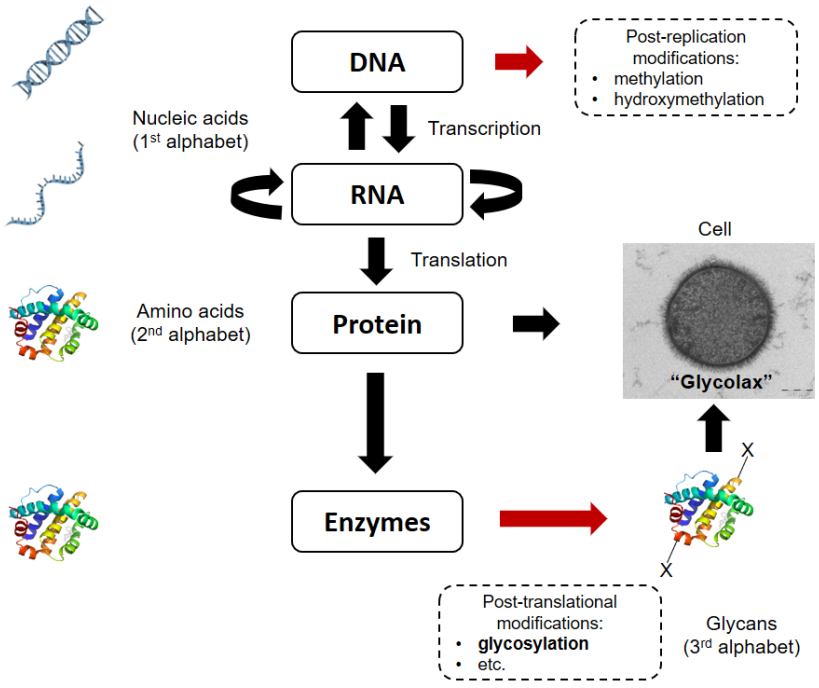

Figure 1. The 20th century central dogma of molecular biology and alphabets of life.

The ubiquitous occurrence of glycans in nature encodes high-density of information ("sugar code"). Comparing the combinatorial complexity of biomolecules, sugars offer an

extremely large coding capacity than nucleic acid and amino acid due to the linkage variation, branching, anomeric isomers and other post-translational modification of the glycan structures (Fig. 2a). The common "letters" that comprise the sugar language with its corresponding structures, symbols and acceptor characteristics are shown in Fig. 2b (4). The enzymatic machinery composed of glycosyltransferase and glycosidases (termed as "writers" and "erasers", respectively) is responsible for the interconversion, elongation and truncation of the "letters" to generate molecular "words". These messages are then translated into biochemical signals triggering respective responses by carbohydrate-specific receptors, called lectins (the "readers").

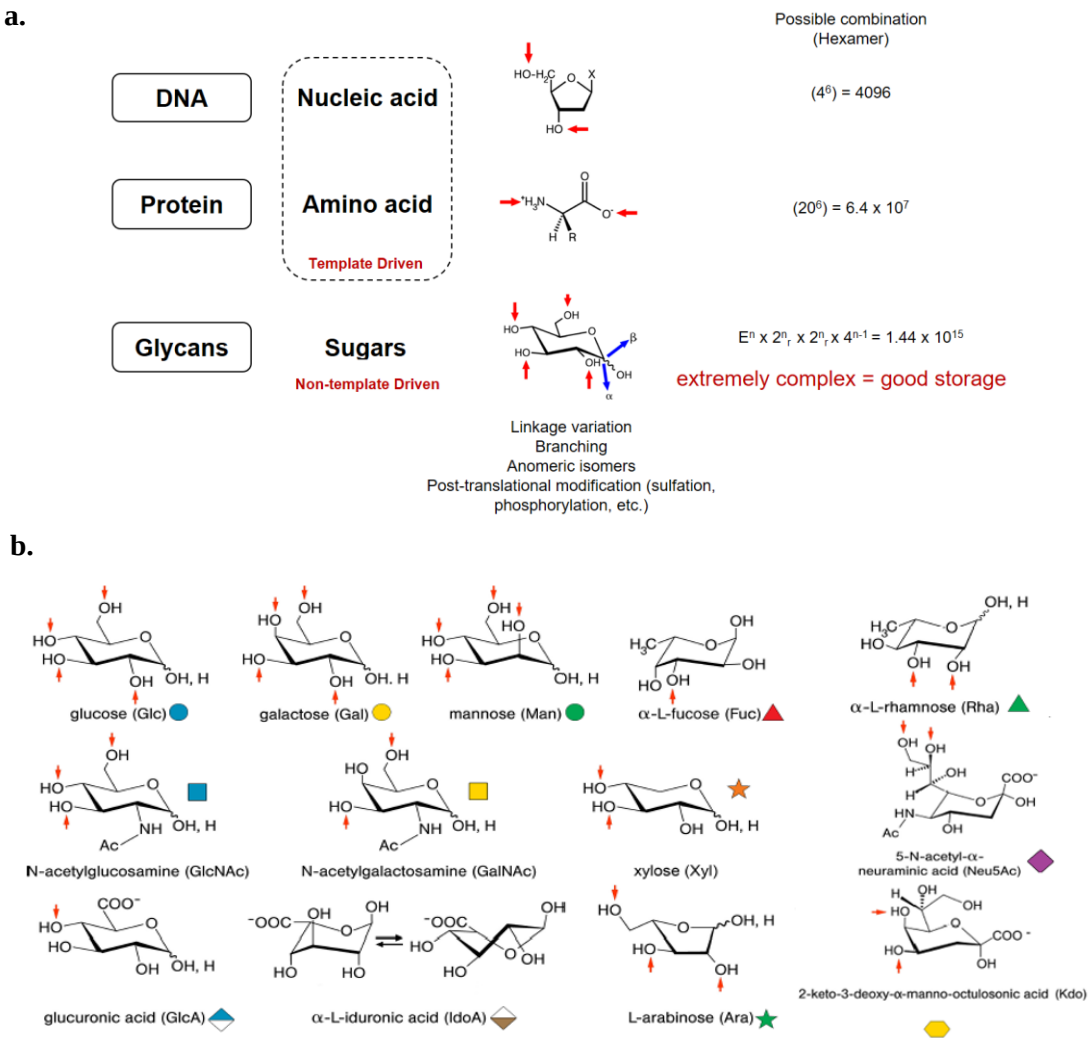

Figure 2. Combinatorial complexity of biomolecules (**A**) and letters of the sugar language (**B**) (4).

Changes in cell surface O-glycosylation patterns (Fig. 3), the most diverse form of protein glycosylation, mediates a wide variety of (patho)physiological functions such as regulation of Notch receptor (by O-fucosylation and O-glucosylation), cell-to-cell signaling, immune response and tumor progression (by mucin (MUC)-type, O-N-acetylgalactosamine glycosylation) (5)(6). Protein O-mannosylation (*O*-Man), an evolutionarily conserved modification in fungi and metazoans. In animals, the O-linked mannose exists in a limited number of proteins that are required for normal development and have vital functions in muscle and neural physiology (7).

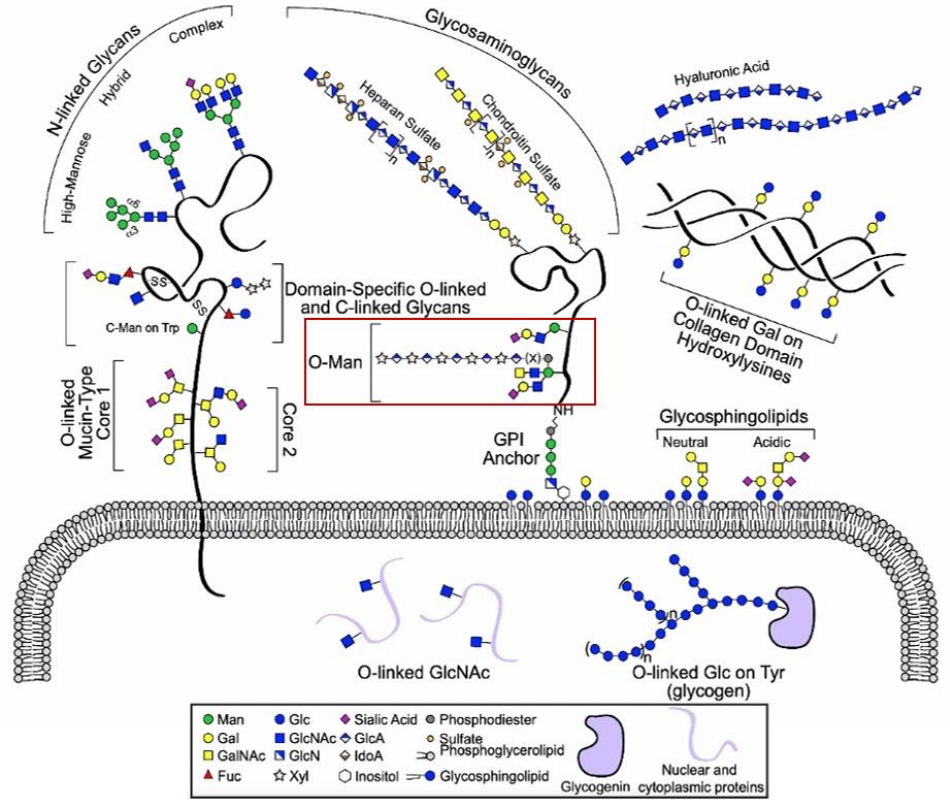

Figure 3. Major classes of vertebrate glycan structures (8).

1.2 α-Dystroglycan

α-dystroglycan (α-DG) is the extracellular component of dystroglycan (DG) (Fig. 4), and is the most extensively studied mammalian *O*-Man glycoprotein. It is ubiquitously expressed in the skeletal muscles and the brain and is associated with cell adhesion, muscle

integrity, neurological development, basement membrane assembly, epithelial polarization, and morphogenesis (9)(10)(11)(12). α-DG also acts as a receptor for pathogens such as *Mycobacterium leprae* and arenaviruses (13)(14). The aberrant glycosylation of α-DG leads to a group of congenital muscular dystrophy (CMD) referred to as α-dystroglycanopathy (α-DGpathy), a neuromuscular disorder which is characterized by progressive muscular degeneration, inaccurate brain formation, intellectual disability, and ocular anomalies (7)(15). These aberrations are currently intractable and various therapeutic approaches have been actively explored (16).

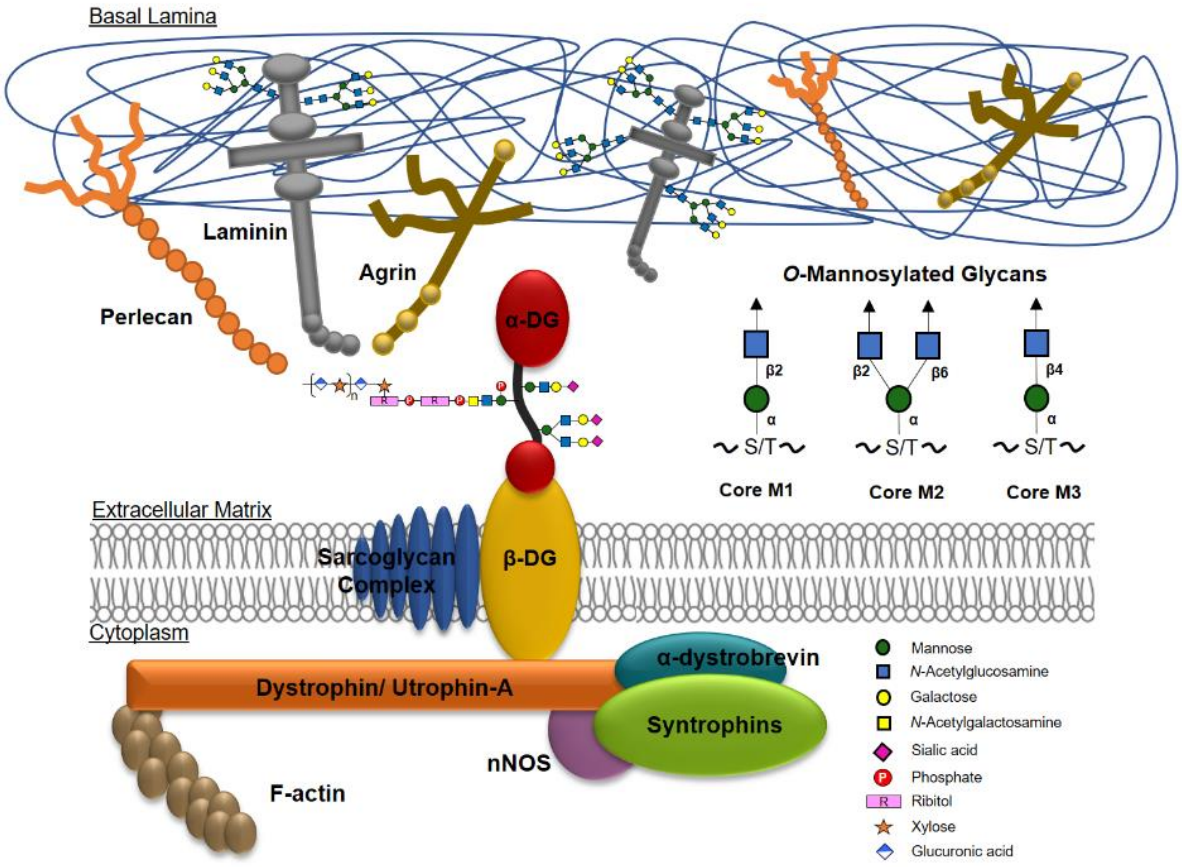

Figure 4. The dystrophin-glycoprotein complex (DGC) presents dystroglycan with its extracellular and cytoskeleton partners. Several types of O-mannosylated glycans identified on α-DG are also depicted.

α-DG possesses unique glycans, i.e. NeuAcα2-3Galβ1-4GlcNAcβ1-2Manα1-O-Ser/Thr, in its mucin (MUC)–like domain (amino acid residues 317–488) (17). Subsequent

studies have identified three types of O-Man glycan structures and have classified them based on the linkage of GlcNAc to Man by different glycosyltransferases: core M1 (GlcNAcβ1-2Manα1-O-Ser/Thr) by protein O-linked mannose β1,2 N-acetylglucosaminyltransferase-1 (POMGNT1); core M2 (GlcNAcβ1-6(GlcNAcβ1-2)Manα1-O-Ser/Thr) by POMGNT1 and then β1,6 N-acetylglucosaminyl transferase-IX/-Vb (GnT-IX/-Vb); and core M3 (GlcNAcβ1-4Manα1-O-Ser/Thr) by POMGNT2 (11), as shown in Fig. 4. The DG in skeletal muscle is intimately involved in cell–extracellular matrix (-ECM) communication via the interaction between the α-DG extended core M3 (matriglycan) and the extracellular carbohydrate-binding proteins (CBPs), such as laminin-1, and -2, and agrin, that contain laminin G-like (LG) domains (11)(18). This association facilitates the interaction of the β-DG with various cytoplasmic proteins, such as dystophin and utrophin, and ultimately with the F-actin, establishing the extra- and intracellular connections. This DG-LG relay system mediates the assembly of the basement membrane, providing muscular stability. Although the functional role of core M3 has been extensively studied, the physiological roles of the core M1-related structures are relatively less investigated.

The role of dystroglycan has been explored mainly in the context of muscles; however, recent studies have also demonstrated the role of DG in the nervous system and have provided additional insights into the brain abnormalities associated with some forms of muscular dystrophies (19). Interestingly, the *O*-Man type glycans constitutes a significant portion of mammalian brain O-glycans (20), with the core M1 and M2 structures accounting for 15% and ~5%, respectively (21). In α-DG, about half of *O*-glycans are *O*-Mannosylated (22). Notably, proper expression of the core M1 glycans is necessary for the core M2 and functional core M3 extension of the α-DG (23). The complete loss of the extended core M3 structures owing to mutations in POMGnT1, fukutin (FKTN), and fukutin-related protein (FKRP) results in typical forms of α-DGpathy, such as muscle-eye-brain disease (MEB), Fukuyama congenital

muscular dystrophy (FCMD), congenital muscular dystrophy type 1C, and limb-girdle muscular dystrophy (7)(24). Intriguingly, a complete loss of MUC-type core 2 *O*-glycans and branched core 1 antennae has been found to lead to a possible functional compensatory upregulation of the *O*-Man glycans in the stomach of engineered mice (25). Since the mentioned MUC-type *O*-glycans are docking sites for tissue lectins underlying a broad range of (patho)physiological activities (26), it is interesting to explore the capacity of *O*-Man glycans to interact with the human lectins, such as the adhesion/growth-regulatory galectins.

1.3 Galectins

Galectins is a family of soluble carbohydrate binding protein (CBP) that share a β-sandwich-type with conserved carbohydrate recognition domain (CRD) that specifically binds to β-galactosides and pair with distinct counterreceptors to trigger a broad range of post-binding events, such as cell growth/adhesion/differentiation, immune response, inflammatory functions, and tumor development and progression (27)(28)(29)(30). Based on the structural aspects of human galectins, three designs can occur in nature: prototype, chimera type and tandem-repeat type (Fig. 5). Each member of this endogenous lectin family (”readers”) deciphers and translates glycan messages depending on its modular architecture as well as the glycan structure and the modification on the galactose residues (“sugar code”) (4)(31). These galectin interactions are physiologically relevant mainly thorugh *cis*-binding and *trans*-bridging of cell surface-associated components in the skeletal muscles and nervous system (Fig. 5) (32)(33)(34).

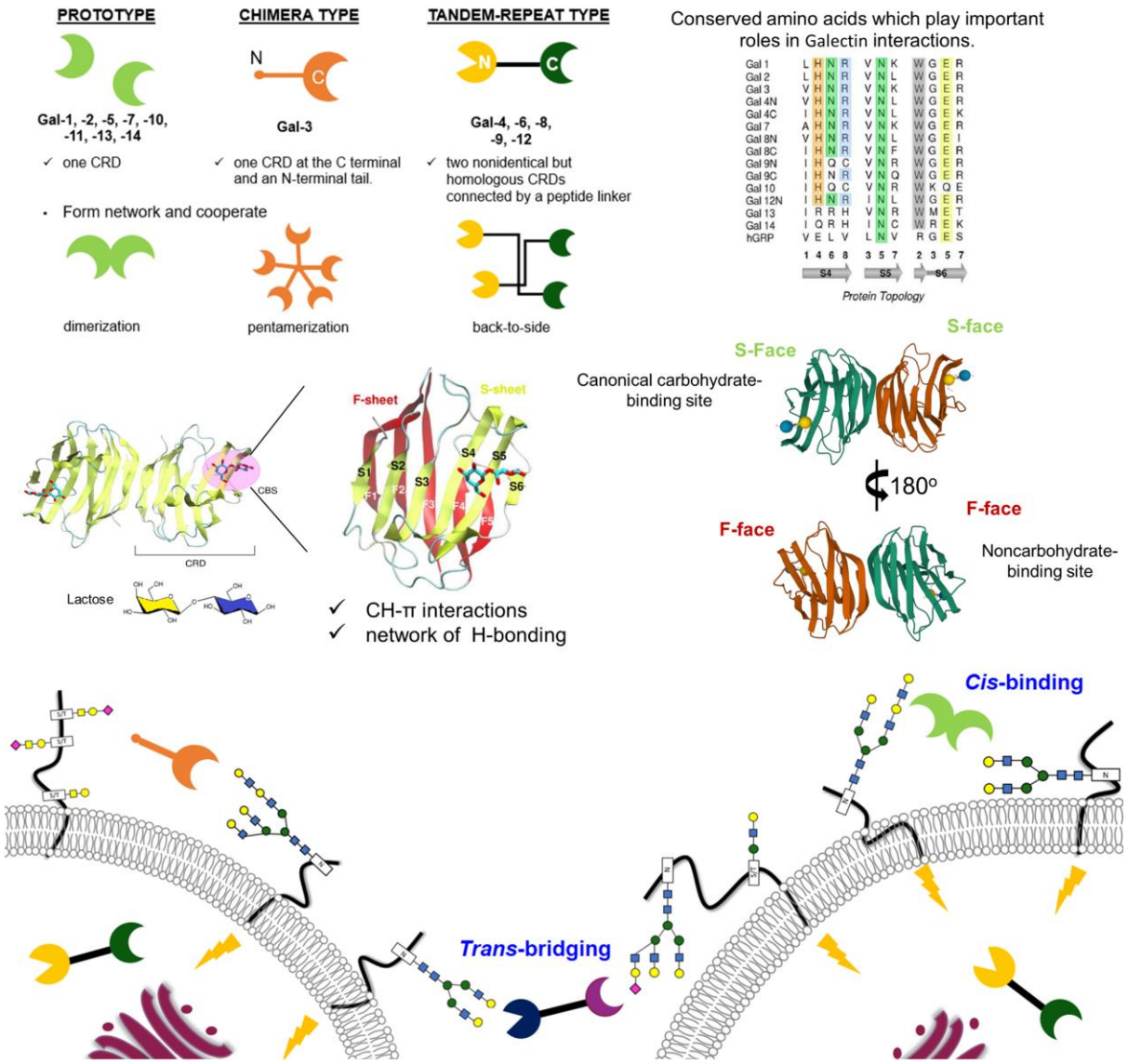

Figure 5. The modular architecture and conserved amino acid sequence across galectin family. CRD positioning relative to the termini is designated by N and C.

Galectins are expressed as cytosolic proteins, and reside in the cytoplasm or nucleus. They exported extracellularly to reach their galactoside ligands after non-classical secretion that by passes the Golgi complex, of which the mechanism remains unknown (35). The galectin CRD is comprised of eleven antiparallel β-strands: S-face [β1 (S1), β10 (S2), β3-6 (S3-S6)] and F-face [β11 (F1), β2 (F2), β7-9 (F3-F5)] (Fig. 5) (27). Lactose (Lac) and *N*-acetyllactosamine (LacNAc) are the simplest sugar to which galectins bind from millimolar to micromolar range depending on the analytical technique utilized and the aglycone moiety, such

as peptide or lipid, to which these glycan units are conjugated (36). For example, the binding of Lac with Gal-1 involves hydrogen bonding (C4-OH, and C6-OH of Gal and the C3-OH of Glc), electrostatic interactions, and van der Waals interactions (CH-π interactions) via stacking of the galactose ring and the highly conserved aromatic ring of tryptophan W68 located on the S-face , as shown in Fig. 5. These interactions with Lac and LacNAc are observed for all galectin CRDs (37). Non-galactose-containing ligands can also interact with galectins via its non-carbohydrate binding sites (F-face) (38)(39). During development, all cells express galectins in varying patterns between cell types and tissues (40). And due to its multivalent oligomerization nature, galectins forms network ("lattice") and cooperate controlling wide-range of biological phenomena (Fig. 6) (41).

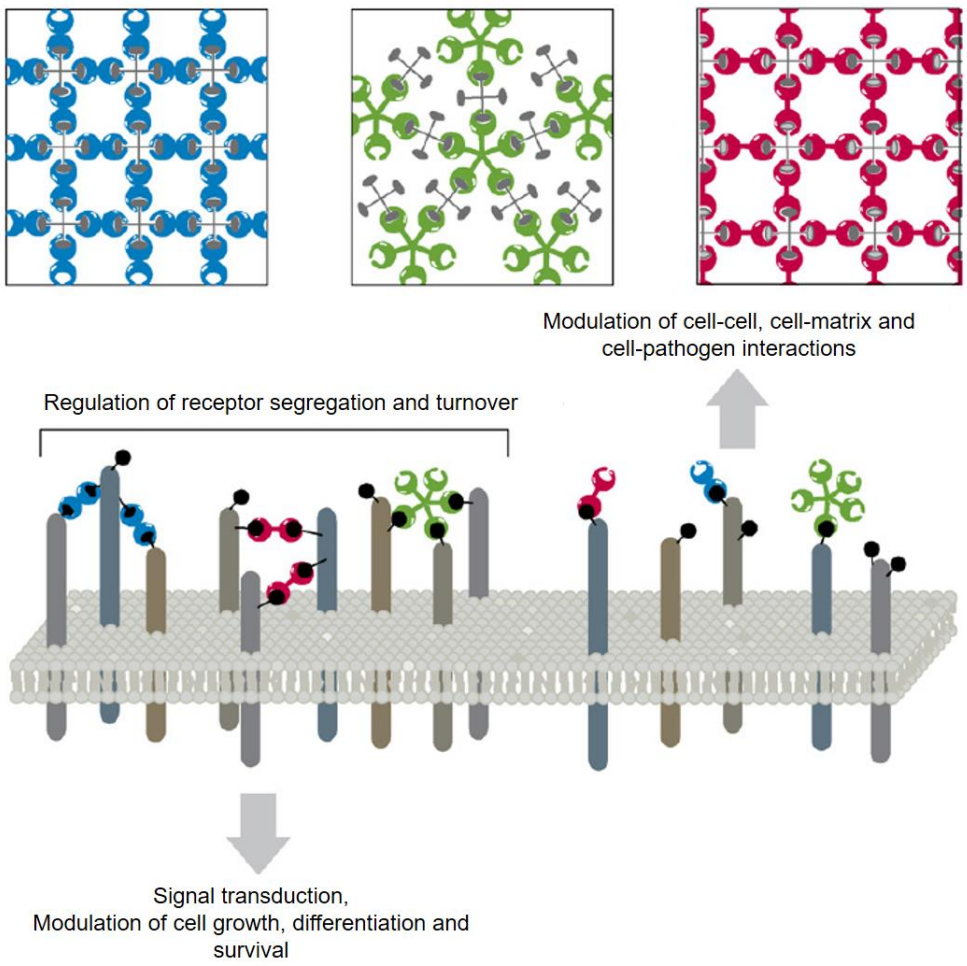

Figure 6. Lattice formation between multivalent galectins and multivalent carbohydrate ligands and its biological relevance (41).

Galectin (Gal)-1, -3, -4 and -9 are expressed by the muscle and neural cells exhibiting critical, yet conflicting roles that remain elusive. A model study has shown that Gal-1, released extracellularly by the differentiating skeletal muscle, interacts with the polyLacNAc[(3Galβ1,4GlcNAcβ1)n]-terminated N-linked glycans of laminin (42). This inhibits the interactions between laminin and α7β1 integrin, thereby regulating myogenesis (32). In addition, the time-dependent expression of α2,3- and α2,6-NeuAc-linked glycans was accompanied by Gal-1 upregulation and Gal-3 downregulation, suggesting a complex interplay between the glycans and galectins during myogenesis (43). In the nervous system, Gal-4 binds to the LacNAc termini of N-glycan structures attached to the neural cell adhesion molecule (NCAM) L1, promoting axonal growth while also regulating oligodendrocyte progenitor cell (OPC) differentiation during the development of nervous system (44). On the contrary, Gal-9 binds to the N-linked carbohydrate motifs of T cell immunoglobulin and mucin domain containing 3 (Tim-3), promoting production of microglial cytokines, such as tumor necrosis factor and interleukin-6 during neuroinflammation (45). These galectin interactions are physiologically relevant mainly through cis-/trans-bridging of the cell surface-associated components in the skeletal muscles and nervous system. Moreover, the viral adhesins on certain viruses, such as coronaviruses, show similarities in the folds that might result in the conformational arrangements favorable for the binding and entry of the virus (46).

1.4 Purpose and strategy

Evidently, the presence of glycan chains on the cellular glycoproteins is functionally relevant when pairing with tissue lectins. This phenomenon has aroused an interest in exploratory studies involving synthetic glycopeptides and lectin panels. Our recent work using chemoenzymatic synthesis and microarray technology has demonstrated that the core M1

extended mucin-like domain of the α-DG fragment, i.e. ^{372}TRGAIIQTPTLGPIQPTRV^{390}, interacts with various plant lectins but not with laminins (47). Using this experimental set-up, glycopeptide microarray technology has been previously shown to be a robust tool for determining the selective pairing of galectins with their binding partners in the MUC1-based glycopeptide library (48). In line with previous works, this study investigated the interaction profile of galectins with the core M1 glycan-bearing glycopeptides. First, the glycoconjugate library was synthesized and their binding to the plant lectins was used as internal controls. The microarray interactions with the wild-type galectin test panel were profiled and the crosslinking capabilities of galectin with the core M1 α-DG glycoconjugates with various laminins were assessed. The nuclear magnetic resonance (NMR) experiments with galectin validated the affinity of galectin to the glycoconjugates. The *trans*-bridging activity of Gal-1 between the *O*-Man core M1 ligands and laminin *in situ* was assessed to propose an additional mechanism by which Gal-1 ameliorates muscular dystrophy. To further delineate the protein design-functionality of this binding process, galectin variants were prepared and their lectin activity towards the prepared ligands were evaluated. The *trans*-bridging capabilities of Gal-1 variants were also determined for possible translational applications of these galectin variants in treatment for some forms of α-dystroglycanopathy.

1.5 References

1. R. Stern, M. J. Jedrzejas, Carbohydrate Polymers at the Center of Life's Origins: The Importance of Molecular Processivity. *Chem. Rev.* **108**, 5061–5085 (2008).

2. A. Varki, Biological roles of glycans. *Glycobiology* **27**, 3–49 (2017).

3. H. J. Gabius, *The sugar code : fundamentals of glycosciences* (Wiley-VCH ; John Wiley [distributor], 2009).

4. H. Kaltner, J. Abad-Rodríguez, A. P. Corfield, J. Kopitz, H.-J. Gabius, The sugar code: letters and vocabulary, writers, editors and readers and biosignificance of functional glycaan–lectin pairing. *Biochem. J.* **476**, 2623–2655 (2019).

5. A. Pandey, N. Niknejad, H. Jafar-Nejad, Multifaceted regulation of Notch signaling by glycosylation. *Glycobiology* **31**, 8–28 (2021).

6. A. Magalhães, H. O. Duarte, C. A. Reis, The role of O-glycosylation in human disease. *Mol. Aspects Med.* **79**, 100964 (2021).

7. T. ENDO, Mammalian *O*-mannosyl glycans: Biochemistry and glycopathology. *Proc. Japan Acad. Ser. B* **95**, 39–51 (2019).

8. A. Corfield, Eukaryotic protein glycosylation: a primer for histochemists and cell biologists. *Histochem. Cell Biol.* **147**, 119–147 (2017).

9. C. M. Dobson, S. J. Hempel, S. H. Stalnaker, R. Stuart, L. Wells, O-Mannosylation and human disease. *Cell. Mol. Life Sci.* **70**, 2849–2857 (2013).

10. S. H. Stalnaker, R. Stuart, L. Wells, Mammalian O-mannosylation: unsolved questions of structure/function. *Curr. Opin. Struct. Biol.* **21**, 603–609 (2011).

11. T. Yoshida-Moriguchi, K. P. Campbell, Matriglycan: a novel polysaccharide that links dystroglycan to the basement membrane. *Glycobiology* **25**, 702–713 (2015).

12. E. Hohenester, Laminin G-like domains: dystroglycan-specific lectins. *Curr. Opin. Struct. Biol.* **56**, 56–63 (2019).

13. A. Rambukkana, *et al.*, Role of α-Dystroglycan as a Schwann Cell Receptor for Mycobacterium leprae *Science (80-.).* **282**, 2076 LP – 2079

(1998).

14. S. Kunz, *et al.*, Posttranslational modification of alpha-dystroglycan, the cellular receptor for arenaviruses, by the glycosyltransferase LARGE is critical for virus binding. *J. Virol.* **79**, 14282–14296 (2005).

15. F. Montanaro, S. Carbonetto, Targeting Dystroglycan in the Brain. *Neuron* **37**, 193–196 (2003).

16. P. M. Van Ry, *et al.*, ECM-Related Myopathies and Muscular Dystrophies: Pros and Cons of Protein Therapies. *Compr. Physiol.* **7**, 1519–1536 (2017).

17. A. Chiba, *et al.*, Structures of sialylated O-linked oligosaccharides of bovine peripheral nerve alpha-dystroglycan. The role of a novel O-mannosyl-type oligosaccharide in the binding of alpha-dystroglycan with laminin. *J. Biol. Chem.* **272**, 2156–2162 (1997).

18. Y.-M. Takako, *et al.*, O-Mannosyl Phosphorylation of Alpha-Dystroglycan Is Required for Laminin Binding. *Science (80-.).* **327**, 88–92 (2010).

19. A. R. Nickolls, C. G. Bönnemann, The roles of dystroglycan in the nervous system: insights from animal models of muscular dystrophy. *Dis. Model. Mech.* **11**, dmm035931 (2018).

20. W. Chai, *et al.*, High prevalence of 2-mono- and 2,6-di-substituted manol-terminating sequences among O-glycans released from brain glycopeptides by reductive alkaline hydrolysis. *Eur. J. Biochem.* **263**, 879–888 (1999).

21. S. H. Stalnaker, *et al.*, Glycomic analyses of mouse models of congenital muscular dystrophy. *J. Biol. Chem.* **286**, 21180–21190 (2011).

22. S. H. Stalnaker, *et al.*, Site Mapping and Characterization of O-Glycan Structures on α-Dystroglycan Isolated from Rabbit Skeletal Muscle*. *J. Biol. Chem.* **285**, 24882–24891 (2010).

23. N. Kuwabara, *et al.*, Carbohydrate-binding domain of the POMGnT1 stem region modulates O-mannosylation sites of α-dystroglycan. *Proc. Natl. Acad. Sci.* **113**, 9280 LP – 9285 (2016).

24. J. Liu, *et al.*, A genetic model for muscle-eye-brain disease in mice lacking protein O-

mannose 1,2-N-acetylglucosaminyltransferase (POMGnT1). *Mech. Dev.* **123**, 228–240 (2006).

25. M. N. Ismail, *et al.*, High-sensitivity O-glycomic analysis of mice deficient in core 2 β1,6-N-acetylglucosaminyltransferases. *Glycobiology* **21**, 82–98 (2011).

26. J. C. Manning, *et al.*, Lectins: a primer for histochemists and cell biologists. *Histochem. Cell Biol.* **147**, 199–222 (2017).

27. HIRABAYASHI, J. ed., Recent topics on galectins. *Trends Glycosci Glycotechnol.* **9**, 1–180 (1997).

28. H. Kaltner, *et al.*, Galectins: their network and roles in immunity/tumor growth control. *Histochem. Cell Biol.* **147**, 239–256 (2017).

29. K. Kasai, Galectins: Quadruple-faced Proteins. *Trends Glycosci. Glycotechnol.* **30**, SE221–SE223 (2018).

30. G. García Caballero, *et al.*, How galectins have become multifunctional proteins. *Histol. Histopathol.* **35**, 509—539 (2020).

31. H.-J. Gabius, J. Roth, An introduction to the sugar code. *Histochem. Cell Biol.* **147**, 111–117 (2017).

32. M. Gu, W. Wang, W. K. Song, D. N. Cooper, S. J. Kaufman, Selective modulation of the interaction of alpha 7 beta 1 integrin with fibronectin and laminin by L-14 lectin during skeletal muscle differentiation. *J. Cell Sci.* **107 (Pt 1**, 175–181 (1994).

33. N. K. Mahanthappa, D. N. Cooper, S. H. Barondes, G. A. Schwarting, Rat olfactory neurons can utilize the endogenous lectin, L-14, in a novel adhesion mechanism. *Development* **120**, 1373–1384 (1994).

34. J. McGraw, *et al.*, Altered primary afferent anatomy and reduced thermal sensitivity in mice lacking galectin-1. *Pain* **114**, 7–18 (2005).

35. C. M. Arthur, M. D. Baruffi, R. D. Cummings, S. R. Stowell, Evolving mechanistic insights into galectin functions. *Methods Mol. Biol.* **1207**, 1–35 (2015).

36. A. Varki, *et al.*, *Essentials of glycobiology, third edition* (2017).

37. T. Shirai, *et al.*, Crystal Structure of a Conger Eel Galectin (Congerin II) at 1.45Å Resolution: Implication for the Accelerated Evolution of a New Ligand-binding Site Following Gene Duplication. *J. Mol. Biol.* **321**, 879–889 (2002).

38. B.-W. Kim, S. Beom Hong, J. Hoe Kim, D. Hoon Kwon, H. Kyu Song, Structural basis for recognition of autophagic receptor NDP52 by the sugar receptor galectin-8. *Nat. Commun.* **4**, 1613 (2013).

39. S. Li, *et al.*, Sterical hindrance promotes selectivity of the autophagy cargo receptor NDP52 for the danger receptor galectin-8 in antibacterial autophagy. *Sci. Signal.* **6**, ra9 (2013).

40. R. Tazhitdinova, A. V Timoshenko, The Emerging Role of Galectins and O-GlcNAc Homeostasis in Processes of Cellular Differentiation. *Cells* **9** (2020).

41. G. A. Rabinovich, M. A. Toscano, S. S. Jackson, G. R. Vasta, Functions of cell surface galectin-glycoprotein lattices. *Curr. Opin. Struct. Biol.* **17**, 513–520 (2007).

42. Q. Zhou, R. D. Cummings, The S-type lectin from calf heart tissue binds selectively to the carbohydrate chains of laminin. *Arch. Biochem. Biophys.* **281**, 27–35 (1990).

43. R. Blazev, *et al.*, Integrated Glycoproteomics Identifies a Role of N-Glycosylation and Galectin-1 on Myogenesis and Muscle Development. *Mol. Cell. Proteomics* **20**, 100030 (2021).

44. S. Velasco, *et al.*, Neuronal Galectin-4 is required for axon growth and for the organization of axonal membrane L1 delivery and clustering. *J. Neurochem.* **125**, 49–62 (2013).

45. A. J. Steelman, J. Li, Astrocyte galectin-9 potentiates microglial TNF secretion. *J. Neuroinflammation* **11**, 144 (2014).

46. M. L. Klein, A. Romero, H. Kaltner, V. Percec, H.-J. Gabius, From Examining the Relationship between (Corona)Viral Adhesins and Galectins to Glyco-Perspectives. *Biophys. J.* (2020) https:/doi.org/https://doi.org/10.1016/j.bpj.2020.11.020.

47. H. Hinou, *et al.*, Synthetic glycopeptides reveal specific binding pattern and conformational change at O-mannosylated position of α-dystroglycan by POMGnT1 catalyzed GlcNAc modification. *Bioorg. Med. Chem.* **27**, 2822–2831 (2019).

48. G. Artigas, H. Hinou, F. Garcia-Martin, H.-J. Gabius, S.-I. Nishimura, Synthetic Mucin-Like Glycopeptides as Versatile Tools to Measure Effects of Glycan Structure/Density/Position on the Interaction with Adhesion/Growth-Regulatory Galectins in Arrays. *Chem. Asian J.* **12**, 159–167 (2017).

Chapter 2.

Microarray screening of galectin and α-dystroglycan core M1 glycopeptide interaction

2.1 Introduction

2.1.1 Glycan microarray platform

Recently, decoding the "sugar code" is gaining much attention to provide deep insight into various pathological processes and designing more effective disease therapeutic strategies and diagnostic tools (1). One of the widely used techniques for high-throughput profiling of glycan-lectin interaction is glycopeptide microarray (2). Our laboratory employs an evanescent-field fluorescence assisted "glycoblotting method" without washing step, a general protocol for the site-specific immobilization of glycopeptides with an aldehyde or an equivalent ketone functional group on slides coated with slide coated with N-protected AO/PC-copolymer (~1.8 nL per spot) under a mild aqueous condition without any coupling reagent (3), Fig. 1. After incubation at 80°C for 1 hr, the printed glycopeptides are covalently attached via stable oxime bond on the microarray slide. The evanescent-field fluorescence-assisted detection system facilitates the analysis of interactions over a broad range of affinities with high reproducibility (4). This technique ensures identical conditions and allows real-time monitoring of the protein-carbohydrate interactions and rapid exchange of the protein analyte without loss of material or surface drying.

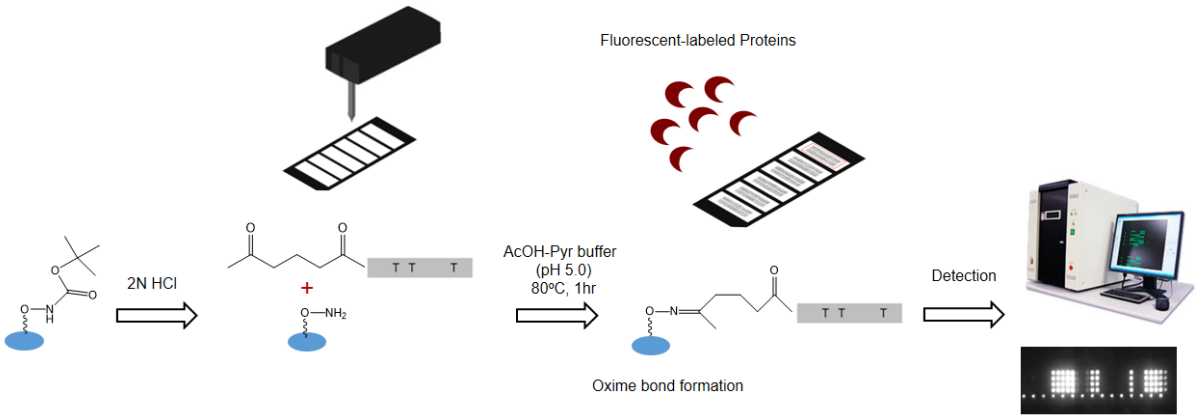

Figure 1. Schematic representation of glycan microarray technology from "glycoblotting method".

The advantage of our technique is demonstrated by highly sensitive and quantitative interaction profiling of MUC1 glycopeptides against galectins and macrophage galactose-type lectin (MGL) (3)(5)(6). Furthermore, the optimized method also works well with analyzing glycoconjugate interaction with fluorophore-tagged antibodies, autoantibodies, plant lectins, and proteins such as lactoferrin and laminins -1 and -2 (7)(8). This method enabled us to determine the precise binding epitope of carbohydrate binding proteins and provide insight into the effect of other glycan moieties, position, and density along the peptide scaffold. Here, we explore the interaction profile of galectins with core M1 glycopeptides of α-DG using glycan microarray technology.

2.2 Results and Discussion

2.2.1 Core M1-based α-DG core M1 glycopeptide microarray and quality control

Our library consisted of 30 compounds comprising the controls and the test glycopeptides as follows: core M1-related glyco-amino acids **1–3**, α-DG peptide backbone **4** (TRGAIIQTPTLGPIQPTRV), glycopeptides **5–29** presenting the core M1-related oligosaccharides at potential *O*-mannosylation sites of α-DG (Thr379, Thr381, and Thr388), and MUC1 peptide **30** (GVTSAPDTRPAPGSTAPPAHGVT) as an additional peptide control (Fig. 2a and SI Appendix, Table S1). Glycoamino acid **1**, nonglycosylated peptides **4** and **30**, and glycopeptides **5–11** with systematically arranged mono-, bis-, and tris-core M1 were synthesized by the standard microwave-assisted solid-phase glycopeptide synthesis methodology terminated with reactive ketone linker at the N-terminus (9)(10). The core M1 disaccharide of compounds **1** and **5–11** was extended by adding a galactose unit using β1,4-galactosyltransferase and UDP-galactose for generating compounds **2** and **12–18**, respectively. These compounds were further elongated by the addition of sialic acid with α2,3-sialyltransferase and CMP-N-acetylneuraminic acid, producing compounds **3** and **19–25**. Glycopeptides **26–29** displaying heterogeneous core M1 structures along α-DG peptide

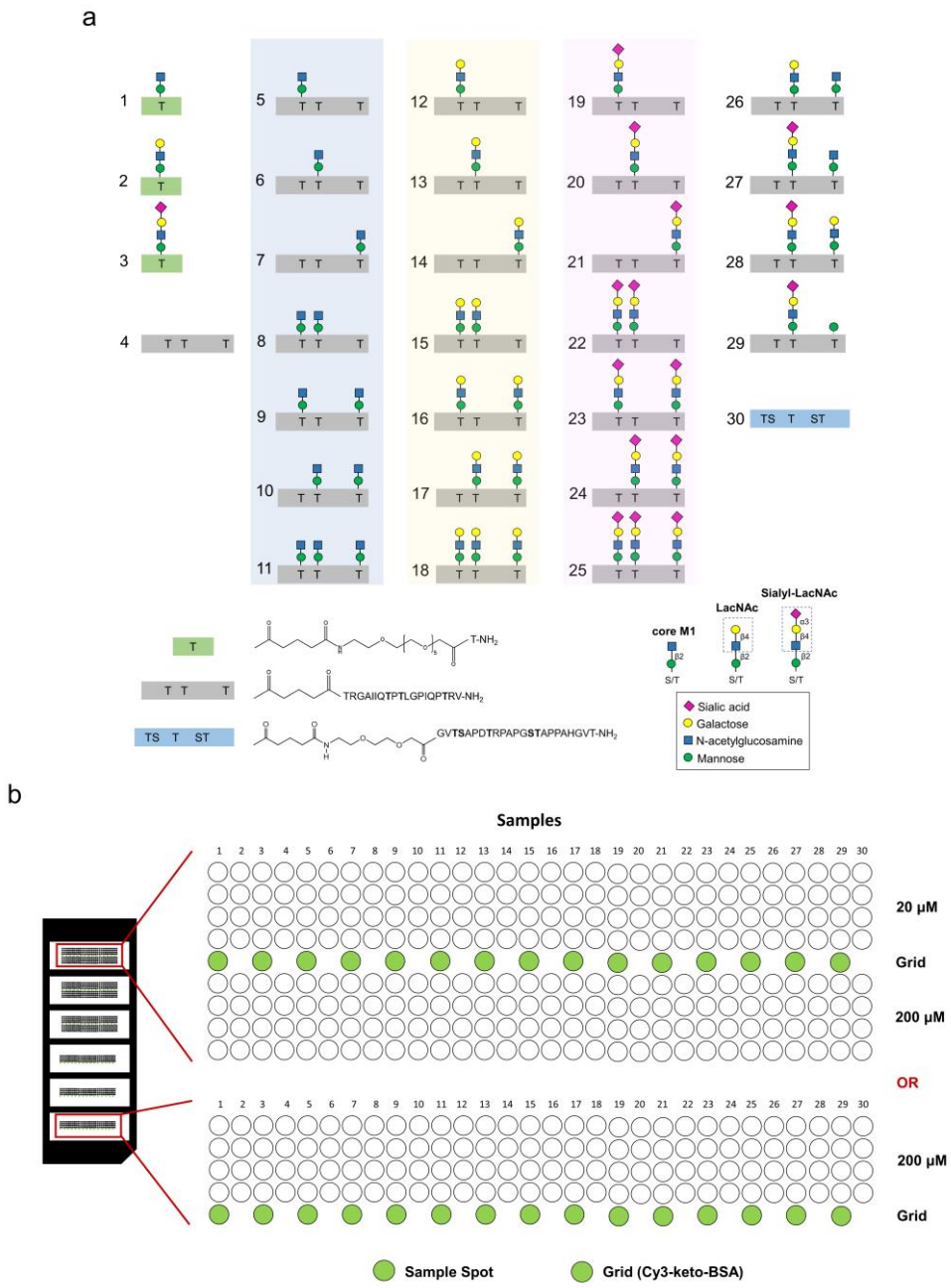

Figure 2. Glycoblotting-based microarray for galectin interaction profiling with core M1 modification. Synthetic α-DG core M1 glycoconjugates used for microarray experiment (**A**). Compounds were robotically printed in quadruplets at 20 and/or 200 µM on an aminooxy- coated plastic slide. Subsequently, a six-chamber rubber silicon sheet was attached. Green spots correspond to cyanine 3-keto-BSA (Cy3-keto-BSA) as grid (**B**).

scaffold were also prepared by a chemoenzymatic protocol. This chemoenzymatically synthesized glycopeptide library can investigate various evaluation at the same time (Fig. 2a): i) the individual contribution of glyco-amino acid and peptide scaffold to the lectin binding (**1-4**) and MUC1 peptide scaffold as negative control (**30**); ii) the impact of individual

glycopeptide and position effect of mono core M1 (GlcNAcβ1-2Man, **5-7**) LacNAc-terminated core M1 (Galβ1-4GlcNAcβ1-2Man, **12-14**) and sialyl-LacNAc-terminated core M1 (NeuAcα2-3Galβ1-4GlcNAcβ1-2Man, **19-21**); and iii) the multivalency effect of bis- and tris-core M1 and extended structures (**8-11**, **15-18**, **22-25**); iv) the effect of different neighboring glycan structure (**26-29**). Subsequently, 20 or 200 μM solutions of compounds **1–30** were used for compound printing onto the surface of a hydroxylamine-functionalized microarray chip in quadruplets, in a six-chamber format (Fig. 2b). For array-based screening with α-DG glycopeptides, the positive and negative control data using plant lectins were presented, before measuring the interaction profiles with a galectin test panel.

Five legume lectins were probed for their accessibility toward cognate glycans, especially mannose, using *Concanavalia ensiformis* agglutinin (ConA) (SI Appendix, Fig. S1). This positive control worked satisfactorily, and the negative controls *Dolichos biflorus* agglutinin (DBA) and *Ulex europaeus* agglutinin I (UEA I) did not generate any carbohydrate-dependent signal, as expected. Galactose or sialic acid were selectively sensed as a minor ligand [*Glycine max* (soybean) agglutinin, SBA or *Triticum vulgaris* (wheat-germ) agglutinin, WGA, respectively] yielded signals at appropriate positions. These results confirmed the validity of the protocol for monitoring the affinity of each member of the galectin panel toward the prepared glycoconjugate library.

2.2.2 Galectin toolbox

Our galectin test panel included three types of architectural design found in humans: the non-covalently-associated homodimeric CRD (prototype) Gal-1; the chimera-type Gal-3 that has a single CRD and a non-lectin N-terminal, inducing self-association forming a pentamer; and the linker-connected heterodimeric CRD with different glycan-binding affinities (tandem-repeat type) Gal-4 and -9 (Fig. 3). These human galectins share binding preference to

Lac/LacNAc, with variability for other types of natural glycans such as LacdiNAc (GalNAcβ1-4GlcNAc) or 3'-O-sulfated galactose, as revealed by systematic frontal affinity chromatography studies (11) or aggregation assays with surface-programmable nanoparticles (12).

Figure 3. Galectin test panel including the wild types of Gal-1, Gal-3, Gal-4, and Gal-9. CRD positioning relative to the termini is designated by N and C.

Among the tested lectins, Gal-1 is unique because it possesses six cysteine residues per CRD/protein making it highly sensitive to oxidation and resulting in decreased or complete loss of lectin activity (13). Thus, previous studies on Gal-1 utilized reducing agents, such as dithiothreitol (DTT), β-mercaptoethanol (βME), reduced glutathione, or iodoacetamide, to prevent oxidative inactivation (14)(15). However, under physiological conditions, the ECM is usually an oxidizing environment (16); hence, we opted to analyze galectin activities in a microarray in the absence of any reducing agents and used ambient air (partially oxidizing conditions) as other galectins. Under non-reducing conditions, Gal-1, but not Gal-3, promotes nerve growth and axonal regeneration at a low concentration (50 pg/mL), the ligand of which is still unknown (17). Moreover, 10 μg/mL Gal-1, -3, and -9 (except for Gal-4) presented almost the same interaction with Galβ1,4GlcNAcβ1,2Manα (LacNAc core M1 glycan) with a thiol linker immobilized in a maleimide-functionalized microarray (18). Therefore, the galectins were tested in a screening concentration range of 0.1 to 32.0 μg/mL in the absence of any reducing agents and used ambient air (partially oxidizing conditions). Since half of the *O*-glycans present in α-DG are mucin-type glycans found in MUC1 (19), before testing the whole core M1 α-DG glycopeptide library, we first surveyed the affinity of galectins to selected

galactose-terminated MUC1 and α-DG glycoconjugates under the fixed microarray conditions (SI Appendix, Fig. S2).

2.2.3 Galactose-terminated MUC1 vs. α-DG core M1 *O*-glycans

The complex *N*-linked and mucin-type *O*-glycan β-galactosides serve as ligands for galectins (20). The highly expressed short determinants of MUC1 glycoprotein in cancer, such as core 1 (TF antigen, Galβ1,3GalNAcα1-O-Ser/Thr) and core 2 (Galβ1-3[GlcNAcβ1-6]GalNAcα1-O-Ser/Thr), were identified as binding receptors for Gal-3, -4, and -8 at 90.0 μg/mL concentration, with no interaction with Gal-1 even at high lectin concentrations (5). Furthermore, a strong interaction of MUC1-bearing TF antigen with Gal-3 is well documented at high concentrations, but not with Gal-1 (21)(22). Herein, the binding of the galectin test panel was preliminarily screened to the selected galactose-terminated MUC1 glycopeptides alongside with the selected core M1 glycoconjugates of α-DG, at 0.1 to 32.0 μg/mL of lectin solutions. The positive controls ConA and PNA (galactose-binding lectin) yielded signals that effectively distinguished between *O*-Man and MUC1 glycoconjugates (SI Appendix, Fig. S3) (8)(23). Notably, only *O*-Man glycopeptides showed interactions at the tested galectin concentrations. This suggests that the tested galectins bind to the *O*-Man ligands with higher affinity than the mucin type *O*-glycans in the array. Thus, the complete screening of the prepared core M1 α-DG glycopeptide library against galectins is beneficial and of interest.

2.2.4 Galectin binding profile with α-DG core M1 glycoconjugates

In a partially oxidizing ambient air environment, the prototype Gal-1 could bind to the glycans of the test panel such that strong signals were recorded (Fig. 4a and b, SI Appendix, Fig. S4 and S5). Gal-1 failed to bind to the core M1 glycoamino acid **1**, whereas β1,4- galactose extension (**2**) led to an enhanced affinity. Further elongation of the core M1 structure with α2,3-NeuAc (**3**) resulted in minimal signal intensity reduction. Notably, binding was observed with unglycosylated peptide **4** of α-DG, but not with MUC1 peptide **30**. In contrast, the core M1 attached to the α-DG peptide backbone (**5–11**) exhibited binding activity to Gal-1. There was

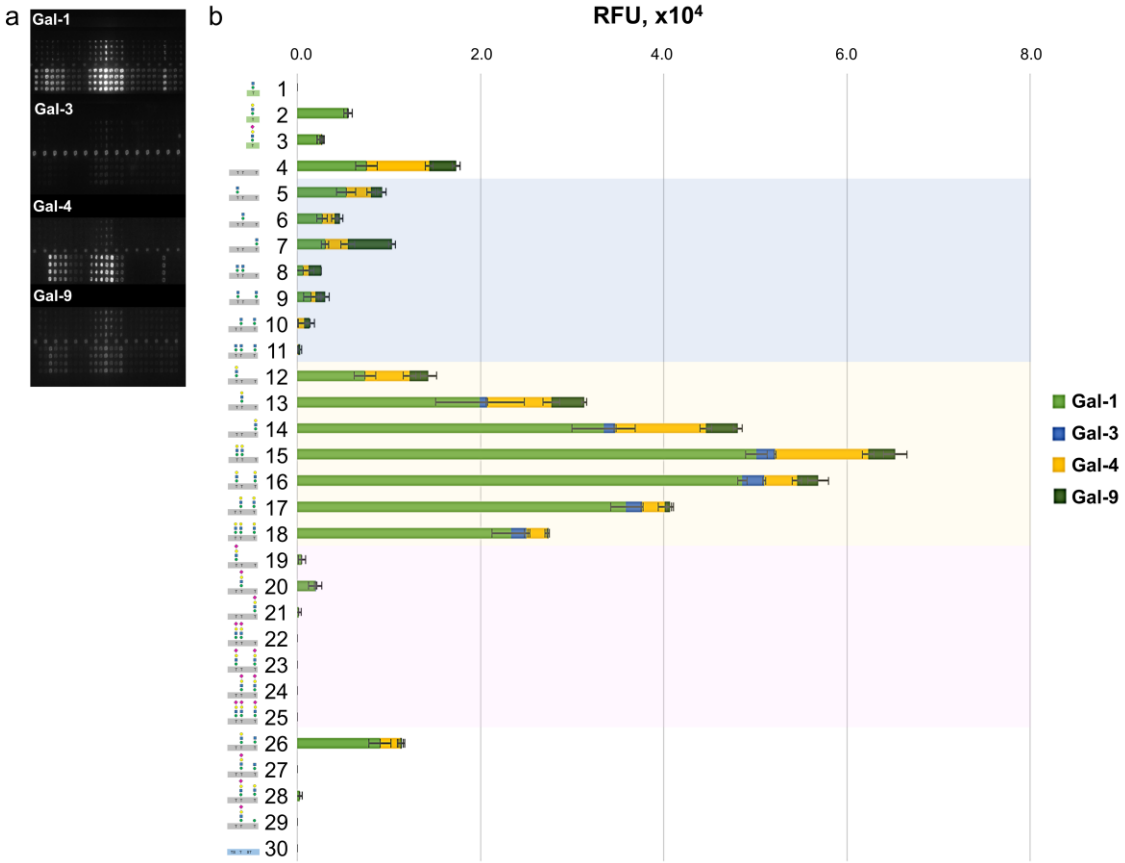

Figure 4. Fluorescence image of microarray chip is taken after treatment of 20 and 200 μM core M1 α-DG glycoconjugates with 10.0 μg/mL galectin solution (a). Stacked chart of signal intensities of 200 μM core M1 α-DG glycoconjugates with 3.20 μg/mL of wild-type galectins (b). The relative interaction of 20 and 200 μM core m1 α-DG glycoconjugates with 0.10 μg/mL to 32.0 μg/mL Gal-1, -3, -4, and -9 are in SI Appendix, Fig. S4 and S5, respectively.

a notable decrease in binding interaction occurred when multiple core M1 disaccharide units were present along the peptide sequence (**5–7** vs. **8–10** vs. **11**). The wild-type protein bound to the LacNAc-terminated core M1 structures (**12–18**, **26**, and **28**). In complex-type biantennary *N*-glycans, Gal-1 has a high affinity for α2,3-sialylated polyLacNAc, while the binding is completely restricted by α2,6-sialylation (24)(25). However, α2,3-sialylation at the LacNAc-terminated core M1 diminished Gal-1 binding to β-galactoside in *O*-Man glycans (**12–18** vs. **19–25**), and even at a single glycosylation site (**17** vs. **28**). This difference in Gal-1 affinity towards α2,3-sialylated N-linked and O-Man glycans is attributed to the conformational change induced by the presence of other carbohydrate moieties along the glycan structure and the type of linkage to the protein (8)(26)(27). Given the redox sensitivity of Gal-1, the binding interaction with α-DG glycoconjugates in a reducing environment containing DTT was verified (SI Appendix, Fig. S6). Under reducing conditions, Gal-1 presented a similar binding behavior but failed to bind to the core M1-terminated 5–11, differentiating Gal-1 activity in the absence of DTT.

The chimera-type Gal-3 showed a weak affinity for the LacNAc-terminated glycopeptides of α-DG (**12–18**) (Fig. 4, SI Appendix, Fig. S4 and S5). This observation was in agreement a recent report where Gal-3 have less interactions with the terminal LacNAc epitopes (28). The position of the LacNAc unit along the peptide sequence (**12** vs. **13** vs. **14**) affected Gal-3 binding. Bis-LacNAc bearing **15** demonstrated stronger affinity than mono-LacNAc **12–14**, because of multivalency at consecutive glycosylation sites. However, longer distance between the two LacNAc units led to reduced interaction (**15** vs. **16–17**). Furthermore, the presence of more than two LacNAc units (**18**) did not result in higher intensity. Similar to Gal-1, these interactions were completely abrogated by the α2,3-sialic acid extension of the LacNAc units (**19–25**).

The tandem-repeat types Gal-4 and -9 presented similar binding patterns but with different signal intensities to α-DG glycoconjugates. Both galectins interacted with peptide **4**, mono-core M1 substituted (**5–7**), and LacNAc-terminated core M1 glycopeptides (**12–18**) (Fig. 4, SI Appendix, Fig. S4 and S5). The influence of LacNAc position and density affected the binding of these lectins as previously observed with Gal-1 and -3. Overall, Gal-4 showed a higher interaction at low concentrations than Gal-9. In contrast, only Gal-9 revealed interactions with glycol-amino acids (**2** and **3**) and α2,3-Sia terminated glycopeptides (**19–25**, **27–29**), thus distinguishing its binding activity with Gal-4. Likewise, α2,3-sialylation of the LacNAc units resulted in a significant reduction in the affinity even with **28** (in the case of Gal-4), which displayed a single LacNAc terminal on one of the glycosylation sites, implying that the different glycosylation of neighboring amino acids can disrupt the affinity of the CBPs.

The glycan-CBP interactions are influenced by not only the type of glycan but also the spatial orientation, carrier scaffold, accessibility, density, and spacing of sugar moieties along the glycoconjugate, to obtain optimum fit within the binding pocket of the receptors (21)(29)(30). In a previous report, Gal-1, -3, and -9, but not Gal-4, revealed similar affinity to the LacNAc- terminated O-Man glycan in a microarray (18). In this study, the addition of Thr to the reducing end of LacNAc-terminated O-Man affected the binding interactions of galectins (Gal-1>>Gal-9>Gal-4, and not Gal-3, Fig. 4, SI Appendix, Fig. S4 and S5). When this LacNAc-terminated core M1 glyco-amino acid structure (**2**) was presented to the α-DG peptide scaffold (**12–14**), the binding affinities of the tested galectins were significantly enhanced. This effect was also observed for the TF antigen–Gal-3 interaction (K_d = 245 μM), which was significantly enhanced when the TF antigen was displayed along the MUC1 peptide scaffold (K_d = 45 μM) (21). Furthermore, bis-LacNAc presenting α-DG glycopeptides demonstrated the highest signal intensity (**15–17**). The tris-LacNAc terminated α-DG glycopeptide (**18**)

showed lesser interaction due to the close proximity of the sugar components along the peptide scaffold resulting in galectin steric hindrance.

2.2.5 Validation of galectin interaction

Gal-1, -4, and -9 (except for Gal-3) were found to interact with peptide **4** of α-DG , but failed to bind with peptide **30** of MUC1, demonstrating the specificity of these galectins

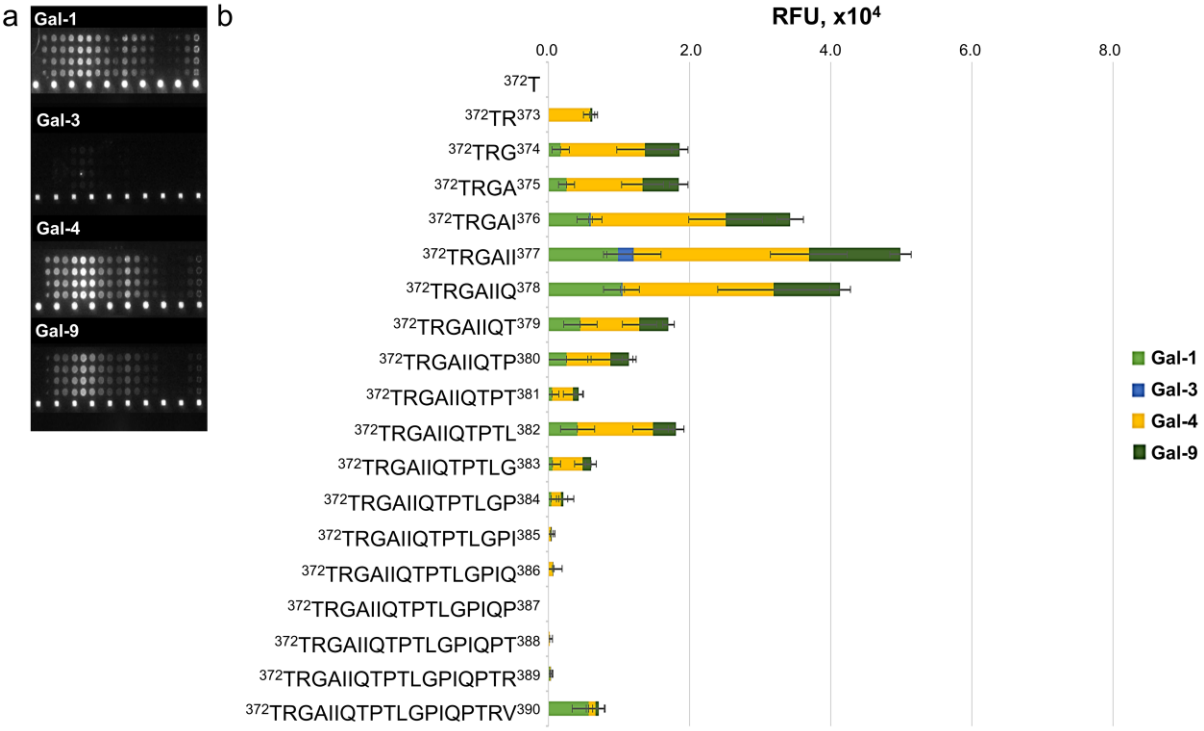

Figure 5. Fluorescence image of α-DG unglycosylated peptide printed microarray chip taken after treatment with 10.0 μg/mL galectin solution (**a**) and relative interaction of 200 μM α-DG unglycosylated peptides with 0.10 μg/mL to 32.0 μg/mL Gal-1, -3, -4, and -9 (**b**). Stacked chart of signal intensities of 200 μM unglycosylated peptide of α-DG library with 3.20 μg/mL Gal-1, -3, -4, and -9 is in SI Appendix, Fig. S7. p = Peptide.

towards certain peptide sequences. Microarray peptide epitope mapping experiments suggested that the presence of hydrophobic amino acids at position ^{374}GAII^{377} is responsible for the binding of unglycosylated peptide **4** to galectins (Fig. 5, SI Appendix, Fig. S7), indicating that specific hydrophobic interactions play an important role. This is analogous to the binding

of Gal-1 to peptide ligands λ5-UR22-45 (52mer, α-helix) and Anginex (33mer, β-sheet) (31)(32). The unglycosylated peptide 4 (19mer, random coil) displays no particular fold similar to the full-length α-DG peptide and needs to be glycosylated to attain a stable structure (33). The presence of core M1 at positions Thr379, Thr381, and Thr388 of the α-DG peptide backbone (**5–7**) lowered the binding affinity. Furthermore, the presence of the bis- and tris-core M1 structures (**8–11**) resulted in a remarkable decrease in the interaction profiles. This can be accounted for by the conformational changes documented in the peptide backbone caused by additional core M1 units along the scaffold, from a random coil-like to a turn-like confirmation (34). This is analogous to our observation on the effect of MUC1 peptide glycosylation and its interaction with macrophage galactose lectin (MGL) (34)(35)(36). Therefore, the glycans attached to the peptide scaffold evidently influence the entropic properties of the peptide backbone by changing the secondary structure (22), thereby affecting the lectin affinity. The galectins revealed high interaction with LacNAc-terminated glycoconjugates (**2**, **12–18**, **26**), further extension of the LacNAc units with α2,3-sialic acid led to a reduction in the galectin affinity, as demonstrated by **3**, **19–25** and **27–29**. To verify whether these observed interactions were dependent on the activity of the CRD, the influence of sugar inhibitor and N-acetylated peptide **4** was determined.

Galectin CRD is comprises of eleven antiparallel β-strands: β1, β10, β3-6 (S-face) and β11, β2, β7-9 (F-face) (37). The presence of methyl-β-lactoside (methyl-Lac), a natural ligand of the canonical sugar-binding site (S-face), resulted in a significantly reduced interaction of galectins with the glycopeptides, specifically the LacNAc-terminated glycoconjugates **2**, **12–18**, and **26** (Fig. 6). Moreover, the addition of non-competing osmolarity controls (maltose and cellobiose) had minimal effect on the binding (SI Appendix, Figs. S8 and S9, respectively). Thus, these binding activities depend on the S-face of the galectins. However, increasing the amounts of N-acetylated peptide **4** in solution was unable to inhibit the affinity of the galectins

to the immobilized ligands (SI Appendix, Fig. S10). Intriguingly, the binding to the compounds **4–18** and **26** (especially **4–7)** was not completely inhibited, when the S-face was blocked by excess (200 mM) methyl-Lac, suggesting that the peptide region of α-DG plays a role in the binding event. These results indicate that galectins can interact with the core M1 glycopeptides of α-DG in two ways: with the LacNAc-terminated sidechains via its S-face and with the peptide backbone via specific hydrophobic interactions.

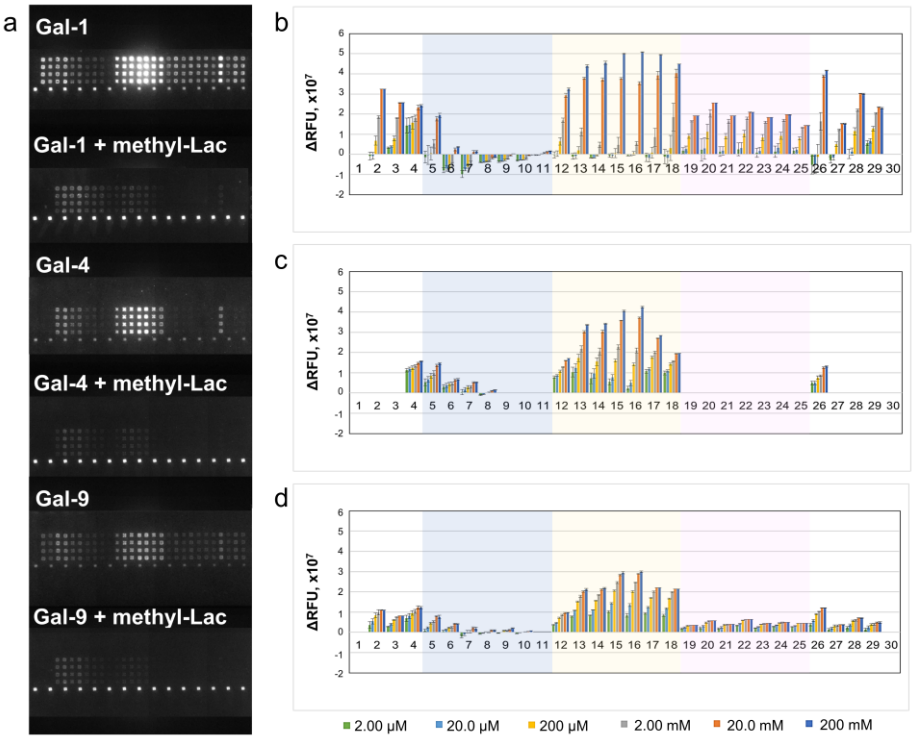

Figure 6. Fluorescence image of core M1 α-DG glycoconjugates printed microarray chip taken after treatment of with 10.0 µg/mL lectin solution and 200 mM methyl-Lac (**a**) and relative binding properties of 200 µM core M1 αDG glycoconjugates with Gal-1 (**b**), Gal-4 (**c**), and Gal-9 (**d**) with the presence of 2.00 µM to 200 mM methyl-Lactose. $\Delta RFU = RFU_{control\ (no\ inhibitor)} - RFU_{methyl\text{-}lac\ (or\ other\ inhibitor)}$

2.3 Conclusion

Combining the availability of the microarray platform with a core M1-type glycopeptide library and a panel of fluorescent galectins, the binding was detected and the profiles were mapped. The type of glycan, position, and density along the α-DG peptide scaffold, as well as the galectin architecture, determined this binding event. Human galectins

can recognize O-Man LacNAc-terminated glycoconjugates, making their respective in situ contacts possible (Gal-1>>Gal-9>Gal-4, and no with Gal-3). The presence of an α2,3-sialylated terminus led to a major reduction in the affinity of galectin, suggesting that this type of extension can fine-tune or act as an on/off switch of galectin activity towards O-Man glycans. These interactions were significantly inhibited by lactose (not by maltose and cellobiose), establishing that the α-DG core M1-type glycans bind to the canonical sugar-binding site (S-face) of galectin, thus serving as a receptor for galectins. The interaction with the peptide region of α-DG, which was not entirely inhibited by lactose, strongly implies that this peptide-galectin interaction occurs via specific hydrophobic interactions that are not dependent on the S-face of galectins.

2.4 Experimental section

2.4.1 Materials

Fmoc-amino acids, Fmoc-TentaGel resins functionalized with Rink amide linker and Sieber Amide linker were purchased from Novabiochem. On the other hand, Fmoc(Ac3GlcNAcβ1→2Ac3Manα)Thr was obtained from Medicinal Chemistry Pharmaceuticals (Sapporo, Japan). Bovine β1,4-galactosyltransferase (β1,4-GalT) and α2,3-sialyltransferase from Pasteurella multocida (α2,3-SiaT) were purchased from Sigma Aldrich. Activated N-acetylneuraminic acid (CMP-NANA) and uridine-5′-diphosphogalactose, disodium salt (UDP-galactose) were acquired from Yamasa Corporation (Chiba, Japan). Other commercially available solvents and reagents were purchased from Sigma-Aldrich (USA), Tokyo Chemical Industry (Tokyo, Japan), Wako Pure Chemical Industries (Osaka, Japan), Kokusan Chemical (Tokyo, Japan), or Watanabe chemical (Osaka, Japan) and used without purification, unless stated.

Manual microwave-assisted solid-phase synthesis was carried out in a polypropylene tube equipped with a filter (LibraTube, Hipep Laboratories, Kyoto, Japan). Green Motif 1 microwave synthesis reactor (IDX Corp, Japan) was used for the glycopeptide synthesis with microwave irradiation. The reaction vessel was placed inside the cavity of the instrument and was irradiated with 2,450 MHz single-mode microwave at 50°C, stirred continuously with a vortex mixer. High-performance liquid chromatography system (HPLC; HITACHI, Japan) was used to purify the compounds using a preparative C-18 reversed- phase column (Intersil ODS-3 10×250 mm) equipped with L7150 pump, at flow rate 5 mL/min monitored by UV detector at 220 nm at room temperature. Bruker Daltonics (Germany) Ultraflex MALDI-TOF/TOF mass spectrometer was used for MALDI experiments using DHB as a matrix.

Microarray fluorescence images were obtained using GlycoStation Reader 1200 (GlycoTechnica Ltd., Yokohama, Japan), analyzed by ArrayVision software V8.0 (GE Healthcare, Tokyo, Japan). Background correction was applied to get the net intensity. The average relative fluorescence unit (RFU) was plotted as a histogram, error bars being the standard deviation utilizing Microsoft Excel.

2.4.2 Methods

Construction of α-dystroglycan mucin type core m1 (glyco)peptide library (5)(8)

The *O*-Man core m1 α-DG glyco-amino acids (**1**), peptide (**4**), glycopeptides (**5-11**), and MUC1 peptide (**30**) were synthesized manually by microwave-assisted solid-phase synthesis by using Fmoc-amino acids (Novabiochem), Fmoc(Ac_3GlcNAcβ1→2Ac_3Manα)Thr (Medicinal Chemistry Pharmaceuticals, Sapporo, Japan), and H-Rink Amide ChemMatrix® (0.48 mmol/g, 24 μmol) resin. The resin was swollen with CH_2Cl_2 in a polypropylene tube equipped with a filter (LibraTube, Hipep Laboratories, Kyoto, Japan) for 1 h at room temperature. The protected Fmoc-amino acid (4.0 equiv) was pre-activated by treating with HBTU (4.0 equiv), HOBt (4.0 equiv), and DIEA (6.0 equiv) in DMF (455 μL) for 9 min under microwave irradiation (Green Motif 1 microwave synthesis reactor, IDX Corp, Japan) and then attached to the resin. In every step, the N-fluorene-9-ylmethoxycarbonyl (Fmoc) groups at N-terminal were removed by 20% piperidine in DMF (1 mL) for 3 min under microwave irradiation. All coupling reactions were done for 10 min, and removed solvents using PP syringes fitted with a porous disk. For glycosylated amino acid, Fmoc-Thr(Ac_3GlcNAcβ1→2Manα1)-OH (1.2 equiv) was treated with PyBOP (1.2 equiv), HOAt (1.2 equiv) and DIEA (3.0 equiv) in DMF (275 μL) subjected to MW irradiation for 9 min at 50°C. After which, PyBOP-HOAt (1.2 equiv) was added and allowed to react for another 9 min. As the final synthesis step, 5-oxohexanoic acid (3 equiv) was introduced at the N-terminus of each

glycopeptide resin, according to the above coupling procedure for Fmoc-amino acids. Removal of side-chain protecting groups and cleavage of glycopeptide from the resin occurred in parallel by treatment with 95% aqueous TFA (1 mL) for 1 h at ambient temperature. The crude peptides and glycopeptides were precipitated in a cold water bath by tert-butyl methyl ether (5 mL). After that, the solution was centrifuged at 3000 rpm for 1 min, and the supernatant was carefully removed. The precipitate was dissolved in milli-Q water (5 mL), and lyophilized. Then, deacetylation of the glycan moiety was done by dissolving the lyophilized material in methanol The pH was adjusted at 12.5 with dropwise addition of 1M NaOH, and the solution was stirred at room temperature for 1 h. After the deprotection, the solution was neutralized with 20% AcOH in methanol, and a flow of nitrogen gas displaced the solvent. The crude peptides and glycopeptides were purified by RP-HPLC, using a preparative C18-reversed-phase column (Intersil ODS-3 10×250 mm) on HPLC (HITACHI, Japan) equipped with L7150 pump, at a flow rate of 5 mL per min monitored by UV detector at 220 nm at room temperature. Eluent A was distilled water containing 0.1% TFA, and eluent B was acetonitrile containing 0.1% TFA. Each product was analyzed by Ultraflex MALDI-TOFMS (Bruker Daltonics, Germany) using DHB as a matrix.

Galactosylation of compounds 1, 5–11 in a 24 h incubation step yielded compounds 2, 12–18. The reaction's 50 mM HEPES buffer (pH 7.0) contains 10 mM $MnCl_2$, 0.1 wt% BSA, galactosyltransferase from bovine milk (Sigma Aldrich), and UDP-Gal (Yamasa Corporation, Chiba, Japan). Subsequently, compounds **2, 12**–**18** were sialylated using α2,3-sialyltransferase from *Pasteurella multocida* (Sigma Aldrich) and CMP-NANA (Yamasa Corporation, Chiba, Japan) in 50 mM Tris buffer (pH 6.5) and 500 mM NaCl incubated for 36 h to yield compounds **3, 19**–**25**. Compounds **26**–**29** were synthesized by solid-phase synthesis and enzymatic sugar elongations, as described earlier. Each product was purified by a RP-HPLC, as described above, with an appropriate solvent system and identified by MALDI-TOFMS.

Preparation and labeling of galectin test panel (38)(39)(40)

Galectins were prepared by recombinant expression using the E.coli strain BL21 (DE) pLysS and the pGEMEX-1 vector (Promega, Walldorf, Germany). Galectin expressing bacteria were grown at 37 °C until an OD600 of 0.6-0.8. Then, the expression of galectin-1 was induced with 100 µM isopropyl ß-D-1-thiogalactopyranoside (IPTG) for at 37 °C, of galectin-3 with 100 µM IPTG at 22 °C, of galectin-4 with 75 µM IPTG at 30 °C and, of galectin-9 with 100 µM IPTG at 22 °C. Induced bacteria were grown for 16 h. Proteins were purified after cell lysis by sonification (three times, each 1 min) through affinity chromatography on a self-made lactose-sepharose resin. Afterwards, bound proteins were eluted from the resin using 20 mM PBS, pH 7,2 containing 50 mM lactose and 20 mM iodoacetamide. PBS was replaced by means of a PD10 column with 10 mM sodium carbonate pH 8.5, and the proteins (2 -3 mg/mL) were directly conjugated in the dark and in the presence of activity-preserving 20 mM lactose, to NHS-ester Alexa 555 fluorescent dye at 25 °C for 4 h. Unbound dye was removed by gel filtration with a Sephadex G-25 column. Protein purity was checked by gel electrophoresis and Western blotting and maintenance of activity by solid-phase, cell and agglutination assays. Labeled proteins were lyophilized in 50 µg aliquots and stored at -20°C until reconstitution. A 1 mg/mL stock probed galectins were prepared in 1x PBS (pH 7.40), containing 1% (w/v) BSA, 0.09% (w/v) NaN_3, and 50% (w/v) glycerol for binding assay and stored at -20°C.

Lectin Binding Assay

The AO/PC-copolymer microarray slides (Sumitomo Bakelite Co., Ltd., Tokyo, Japan) were deprotected using 2N HCl overnight at rt, rinsed with MilliQ H_2O, and dried by centrifugation. The test compounds were robotically printed in quadruplets at two concentrations (20 and 200 µM) in 25 mM AcOH-Pyr (pH 5.0), 0.0025% (w/v) Triton X-100

using an Arduino-based CNC machine handcrafted robot. A cyanin3-keto-BSA (Cy3-keto-BSA) at 25 µg/mL was also printed on the slide as a grid. Subsequently, the slide was incubated at 80°C for 1 h to complete the oxime bond formation. Washed once with Milli-Q H_2O and dried by centrifugation at 2,000 rpm for 2 min.

A silicon rubber sheet with six chambers was attached to the printed slide. Next, slide was pretreated with reaction buffer {Phosphate-Buffered Saline solution (PBS, 1X) [10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2.7 mM KCl, 137 mM NaCl] pH7.4 containing 0.05% (v/v) Tween-20} for 15 min and dried by centrifugation. The plant lectin or galectin solutions in PBS were prepared and maintained in a cold ice bath before use. A cover glass was then set on each chamber, and 12 µL of 0.10 µg/mL lectin solution in reaction buffer was added through the gap of the slide and cover. After 1 h of incubation with the lectin solution at rt in a humidified chamber, the slide was rinsed with washing buffer, and fluorescence intensity was measured with GlycoStation System (GlycoStation Reader 1200, GlycoTechnica Ltd., Yokohama, Japan). To determine additional interactions at higher Gal concentrations, the reaction buffer solution was carefully removed and replaced with the next lectin test concentration, incubated for 5 min at rt, washed with reaction buffer, and obtained fluorescence intensity. This step was repeated until all the chosen test concentrations were completed (0.32-, 1.00-, 3.20-, 10.0-, and 32.0 µg mL^{-1}).

Images of slides were captured in the presence of reaction buffer. The fluorescence intensities obtained were analyzed using ArrayVision software V8.0 (GE Healthcare, Tokyo, Japan). The background correction was applied to get the net intensity. The average relative fluorescence unit (RFU) was plotted as a bar graph and error bars being the standard deviation utilizing Microsoft Excel.

Inhibition experiment

A 12 µL of 10 µg/mL galectin with 2.00 µM of sugar inhibitors (methyl-β-lactoside, cellobiose, maltose) or 100 µM acetylated peptide 4 in PBS was added to the printed test compounds. After 30 min of incubation with the galectin-inhibitor solution at rt in a humidified chamber, the slide was rinsed with washing buffer and fluorescence intensity was measured with GlycoStation System. To determine the effect of higher concentrations of inhibitors, the reaction buffer solution was carefully removed and replaced with the next galectin-inhibitor solution test concentration, incubated for 5 min at rt, washed with reaction buffer and then fluorescence intensity was obtained. This step was repeated until all the chosen inhibitor test concentrations were completed: 20.0 µM, 200 µM, 2.00 mM, 20 mM, and 200 mM for sugar inhibitors; and 200 µM, 500 µM, 1.00 mM, 2.00 mM and 5.00 mM for acetylated peptide 4. Slide images were captured and analyzed as described above.

2.4.3 Supplementary information

Table S1. *O*-Man core m1 based glycopeptide sequence utilized in this study (glycosylated positions in **red** and **blue**).

Code	Glycan	Sequence
1	GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-PEG-**T**-NH_2
2	Galβ(1→4)GlcNAcβ(1→2) Manα1→	5-oxo-hexanoyl-PEG-**T**-NH_2
3	Siaα(2→3)Galβ(1→4) GlcNAcβ(1→2) Manα1→	5-oxo-hexanoyl-PEG-**T**-NH_2
4	None	5-oxo-hexanoyl-TRGAIIQTPTLGPIQPTRV -NH_2
5	GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**PTLGPIQPTRV -NH_2
6	GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTP**T**LGPIQPTRV -NH_2
7	GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTPTLGPIQP**T**RV -NH_2
8	GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**P**T**LGPIQPTRV -NH_2
9	GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**PTLGPIQP**T**RV -NH_2
10	GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTP**T**LGPIQP**T**RV -NH_2
11	GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**P**T**LGPIQP**T**RV -NH_2
12	Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**PTLGPIQPT RV-NH_2
13	Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTP**T**LGPIQPT RV -NH_2
14	Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTPTLGPIQP**T**RV -NH_2
15	Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**P**T**LGPIQPTRV -NH_2
16	Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**PT LGPIQP**T**RV -NH_2
17	Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTP**T**LGPIQP**T**RV -NH_2
18	Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**P**T**LGPIQP**T**RV -NH_2
19	Siaα(2→3)Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**PTLGPIQPTRV -NH_2
20	Siaα(2→3)Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTP**T**LGPIQPTRV -NH_2
21	Siaα(2→3)Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTPTLGPIQP**T**RV -NH_2
22	Siaα(2→3)Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**P**T**LGPIQPTRV -NH_2
23	Siaα(2→3)Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**PTLGPIQP**T**RV -NH_2
24	Siaα(2→3)Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTP**T**LGPIQP**T**RV -NH_2
25	Siaα(2→3)Galβ(1→4)GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQ**T**P**T**LGPIQP**T**RV -NH_2
26	Galβ(1→4)GlcNAcβ(1→2) Manα1→ GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTP**T**LGPIQP**T**RV -NH_2
27	Siaα(2→3)Galβ(1→4)GlcNAcβ(1→2)Manα1→ GlcNAcβ(1→2)Manα1→	5-oxo-hexanoyl-TRGAIIQTP**T**LGPIQP**T**RV -NH_2
28	Siaα(2→3)Galβ(1→4)GlcNAcβ(1→2)Manα1→ Galβ(1→4)GlcNAcβ(1→2) Manα1→	5-oxo-hexanoyl-TRGAIIQTP**T**LGPIQP**T**RV -NH_2
29	Siaα(2→3)Galβ(1→4)GlcNAcβ(1→2)Manα1→ Manα1→	5-oxo-hexanoyl-TRGAIIQTP**T**LGPIQP**T**RV -NH_2
30	None	5-oxo-hexanoyl-PEG-GVTSAPDTRPAPGSTAPPAHGVT-NH_2

Synthesis and characterization of compounds used in this study is reported elsewhere (8).

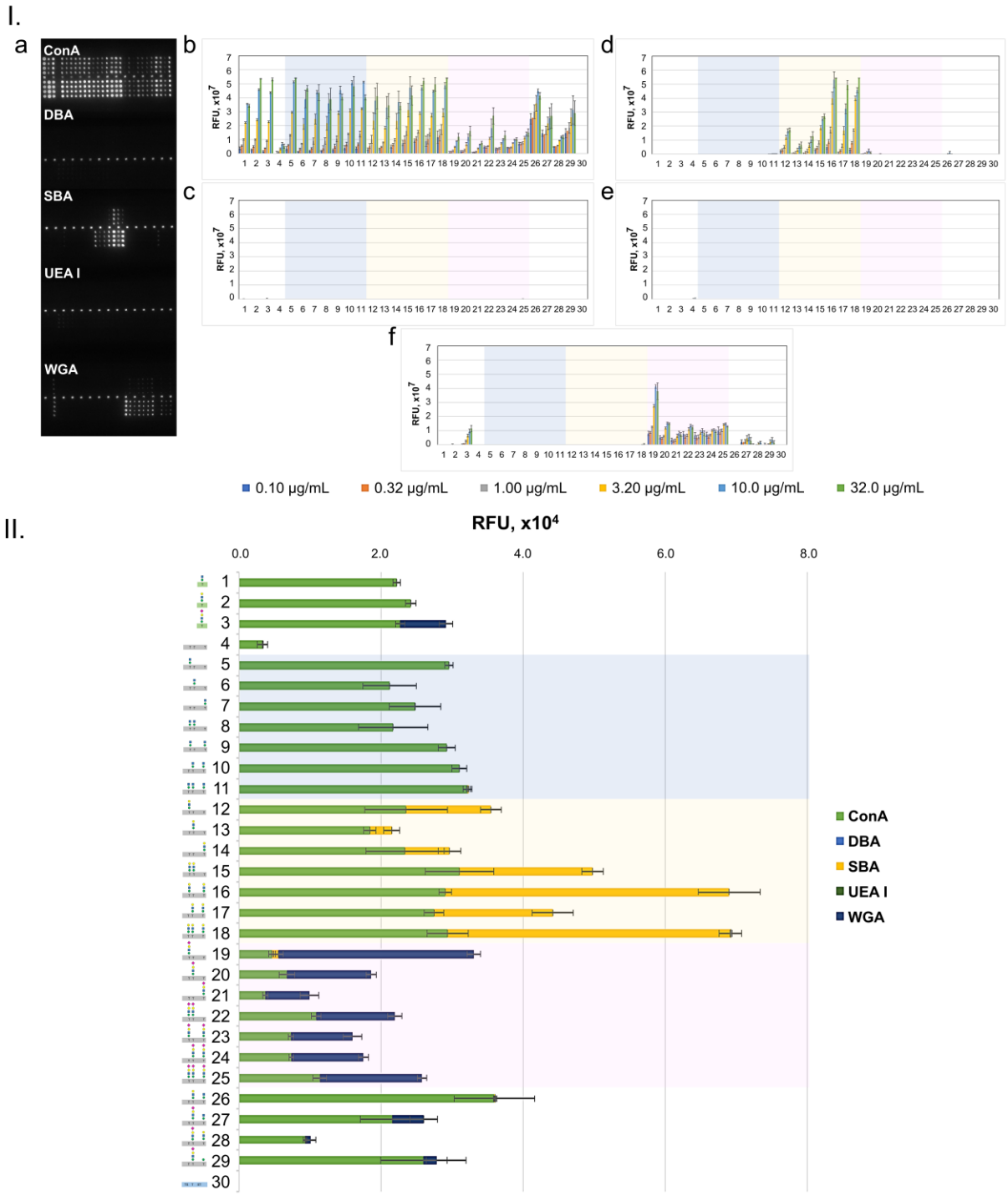

Figure S1. (**I.**) Fluorescence image of microarray chip is taken after treatment of 20 and 200 µM core M1 α-DG glycopeptides with 10.0 µg/mL plant lectin solution (**a**) and relative binding properties of 200 µM core M1 of αDG with 0.10 to 32.0 µg/mL rhodamine-label plant lectins; *Concanavalia enisformis* agglutinin (ConA) (**b**), *Dolichos biflorus* agglutinin (DBA) (**c**), *Glycine max* (Soybean) agglutinin (SBA) (**d**), *Ulex europaeus* agglutinin I (UEA I) (**e**), *Triticum vulgaris* (Wheat germ) agglutinin (WGA) (**f**). (**II.**)Stacked chart of signal intensities of 200 µM core M1 α-DG glycopeptides with 3.20 µg/mL of plant lectins *Concanavalia ensiformis* agglutinin (ConA), *Dolichos biflorus* agglutinin (DBA) , *Glycine max* (Soybean) agglutinin (SBA), *Ulex europaeus* agglutinin I (UEA I), and *Triticum vulgaris* (Wheat germ) agglutinin (WGA).

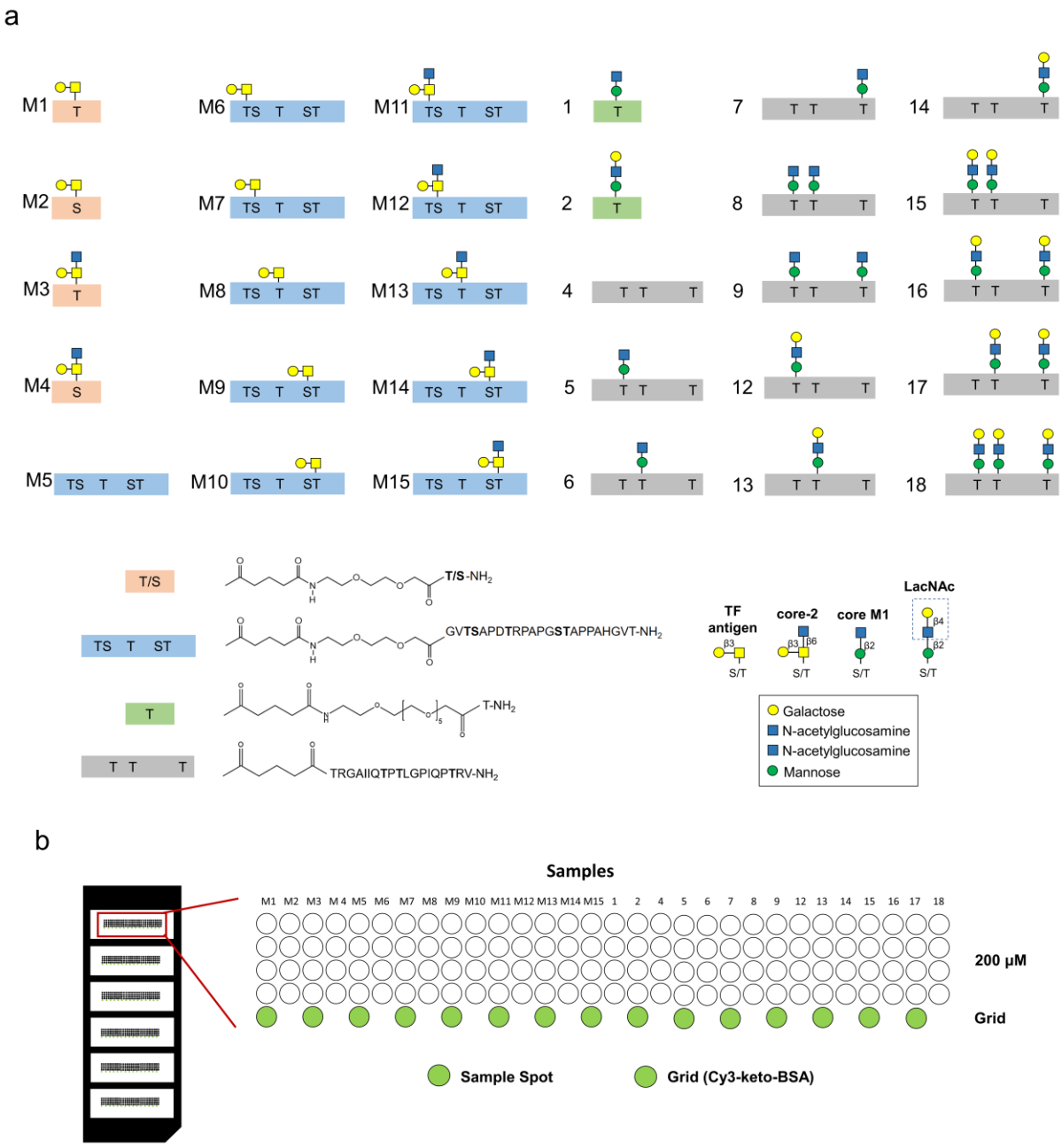

Figure S2. Synthetic MUC1 and α-DG core M1 glycopeptides used for preliminary microarray experiments (**a**). Compounds were robotically printed in quadruplets at 200 µM on an aminooxy- coated plastic slide. Subsequently, a six-chamber rubber silicon sheet was attached. Green spots correspond to cyanine 3-keto-BSA (Cy3-keto-BSA) as grid (**b**).

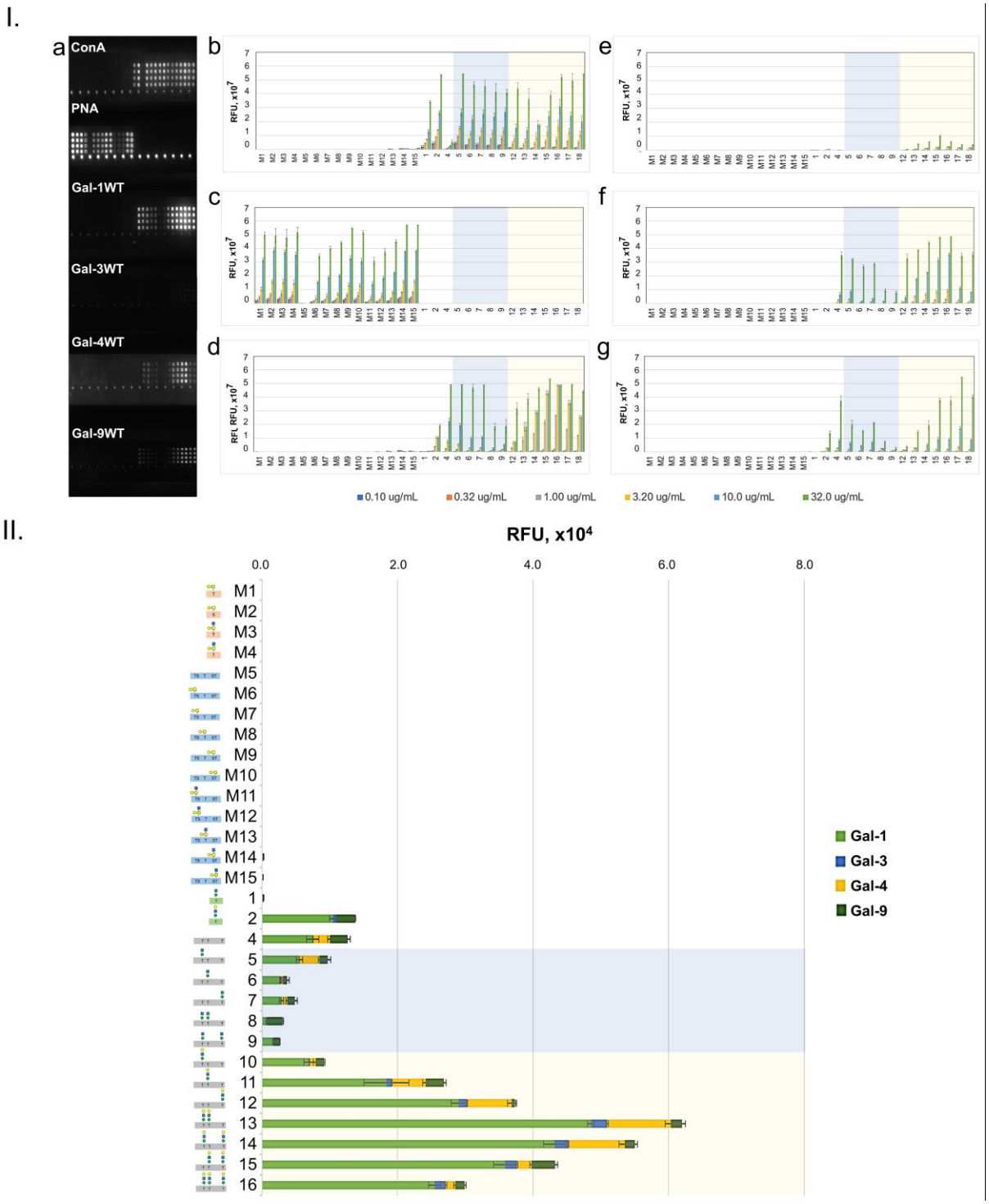

Figure S3. Synthetic MUC1 and α-DG core M1 glycopeptides used for preliminary microarray experiments. (**I**.) Fluorescence image of microarray chip of selected galactose-terminated MUC1 and α-DG glycopeptide peptide library taken after treatment with 10.0 μg/mL galectin solution (**a**) and relative interaction of 200 μM glycopeptides with 0.10 μg/mL to 32.0 μg/mL ConA (**b**), PNA (**c**), Gal-1 (**d**), Gal-3 (**e**), Gal-4 (**f**), and Gal-9 (**g**). (**II**.) Stacked chart of signal intensities of 200 μM of selected MUC1 and core M1 α-DG glycopeptides with 3.20 μg/mL of galectins. The synthesis and characterization of MUC1 glycopeptides used in this study is reported elsewhere[2,3].

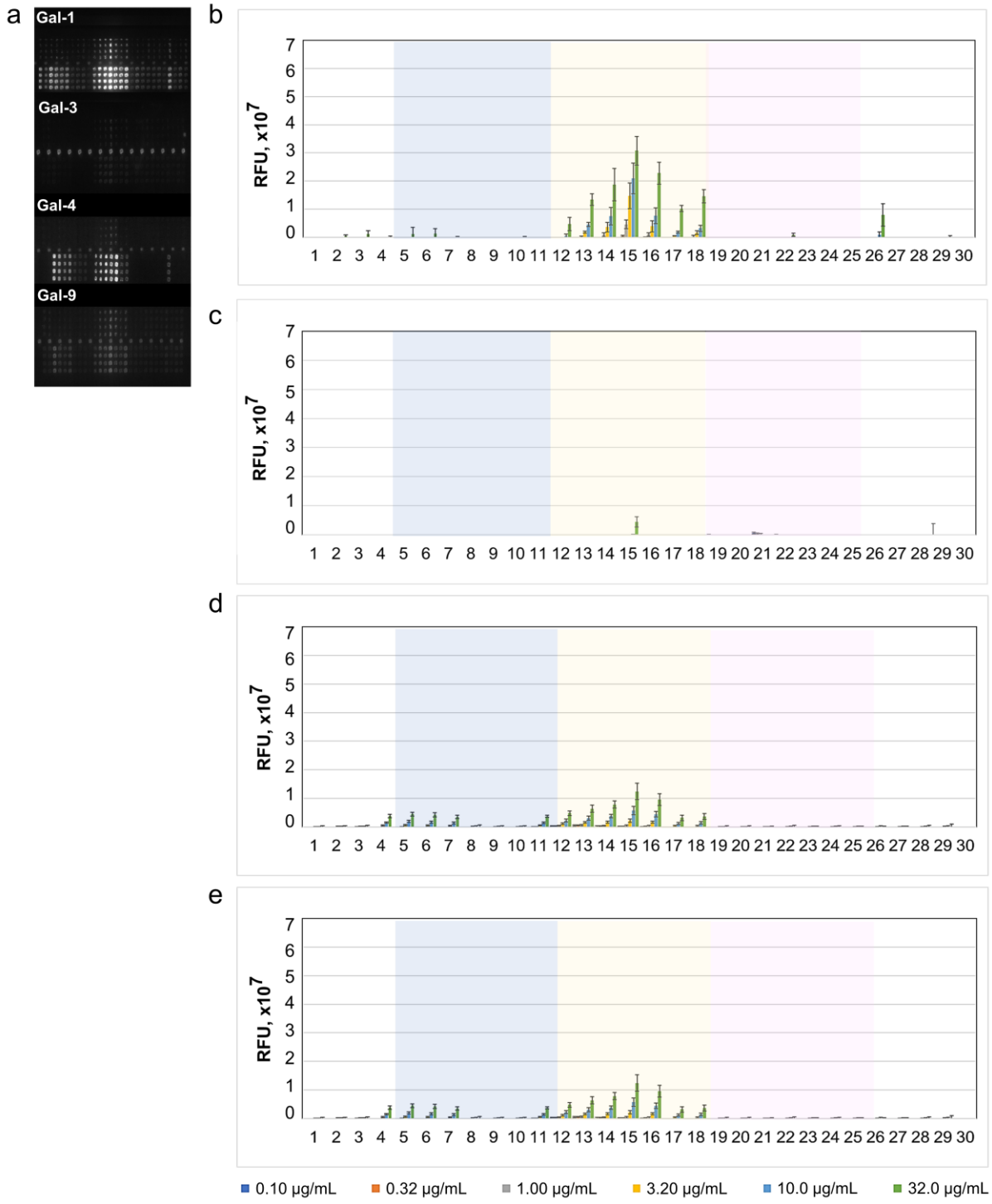

Figure S4. Fluorescence image of microarray chip is taken after treatment of 20 and 200 µM core M1 α-DG glycoconjugates with 10.0 µg/mL galectin solution (**a**) and relative interaction of **20 µM** core m1 α-DG glycoconjugates with 0.10 µg/mL to 32.0 µg/mL Gal-1 (**b**), Gal-3 (**c**), Gal-4 (**d**), and Gal-9 (**e**).

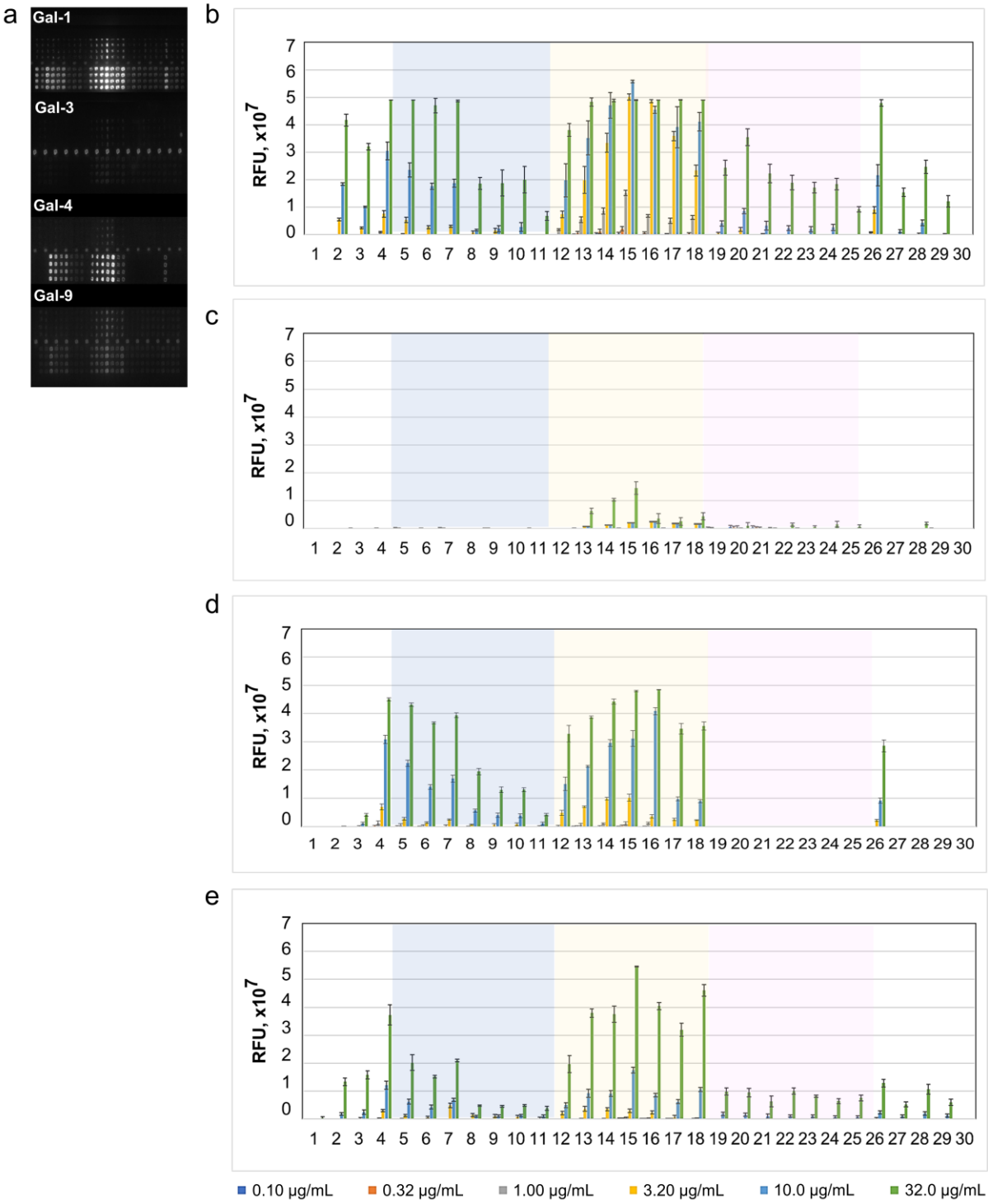

Figure S5. Fluorescence image of microarray chip is taken after treatment of 20 and 200 µM core M1 α-DG glycoconjugates with 10.0 µg/mL galectin solution (**a**) and relative interaction of **200 µM** core m1 α-DG glycoconjugates with 0.10 µg/mL to 32.0 µg/mL Gal-1 (**b**), Gal-3 (**c**), Gal-4 (**d**), and Gal-9 (**e**).

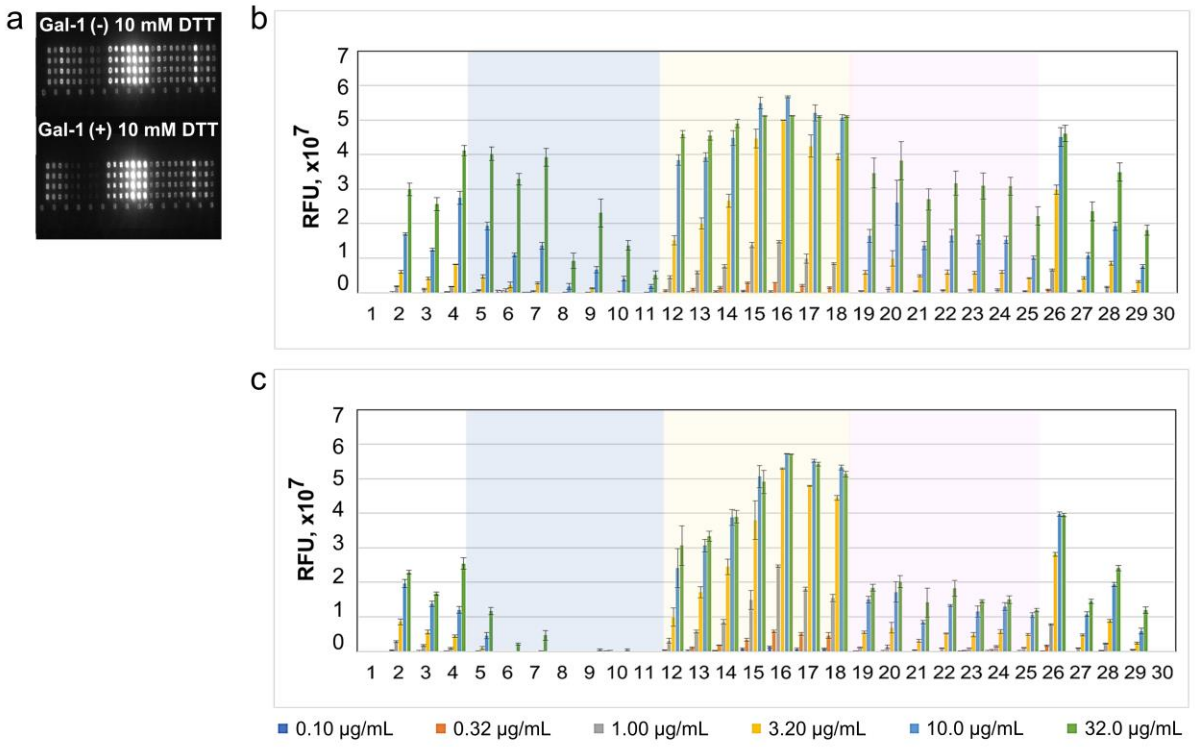

Figure S6. Fluorescence image of microarray chip is taken after treatment of 200 µM core M1 α-DG glycopeptides with 10.0 µg/mL Gal-1 in the absence and presence of 10 mM DTT (**a**) and relative interaction of 200 µM core M1 α-DG glycopeptides with 0.10 µg/mL to 32.0 µg/mL Gal-1 absence (**b**) and presence (**c**) of 10 mM DTT.

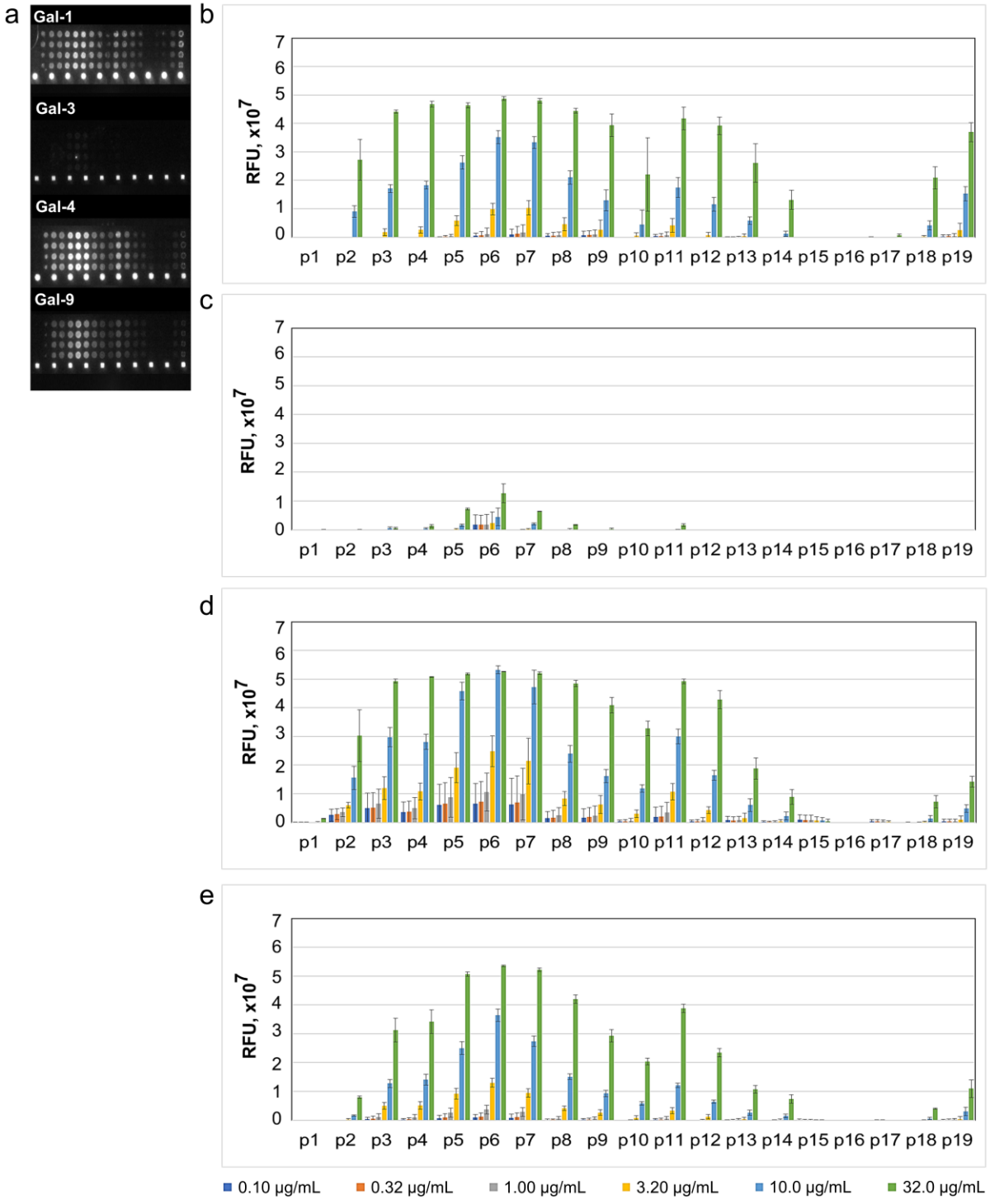

Figure S7. Fluorescence image of microarray chip of α-DG unglycosylated peptide library taken after treatment with 10.0 µg/mL galectin solution (**a**) and relative interaction of 200 µM α-DG unglycosylated peptides with 0.10 µg/mL to 32.0 µg/mL Gal-1 (**b**), Gal-3 (**c**), Gal-4 (**d**), and Gal-9 (**e**).

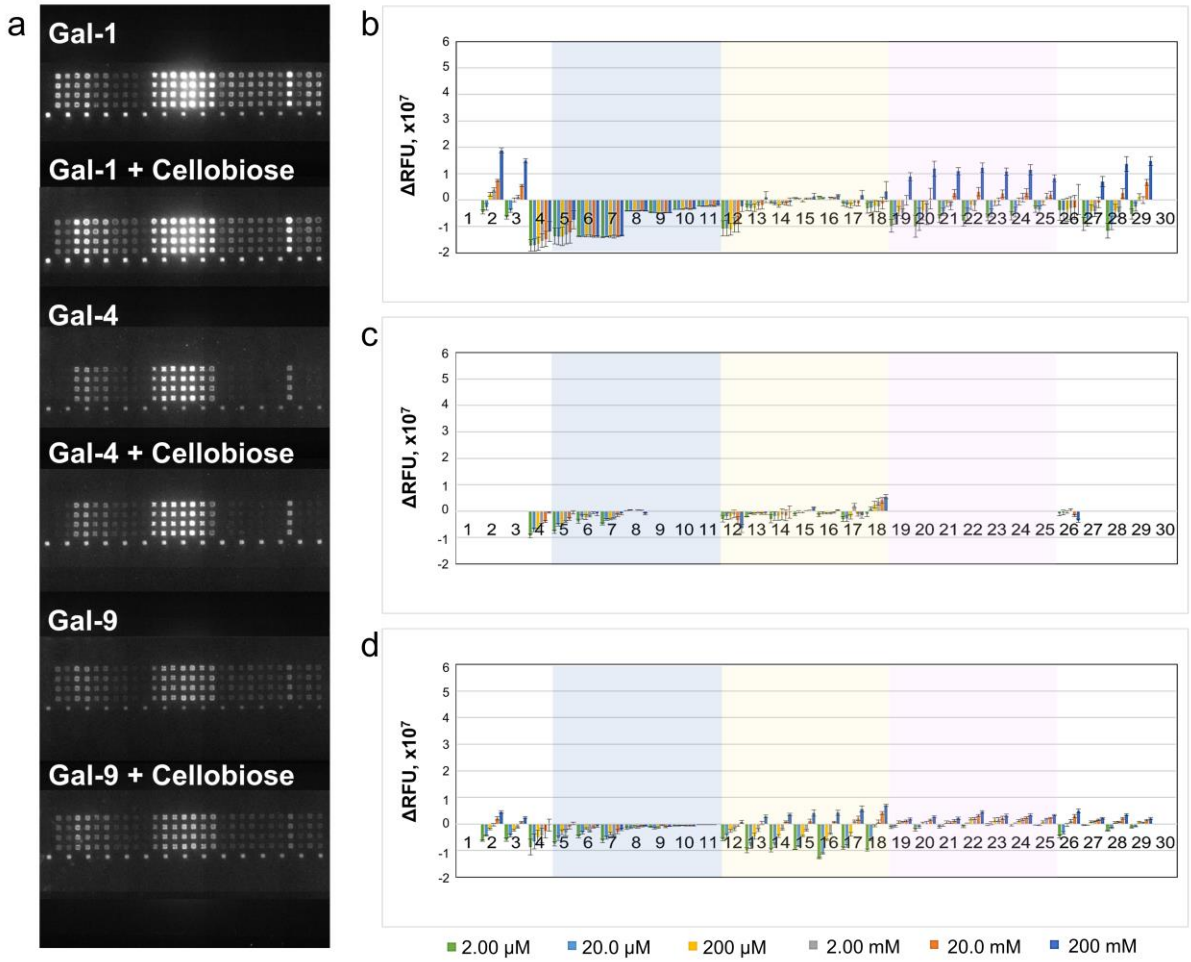

Figure S8. Fluorescence image of microarray chip is taken after treatment of 10 μg/mL lectin solution with 200 mM cellobiose (**a**) and relative binding properties of 200 μM core M1 α-DG glycopeptides with Gal-1 (**b**), Gal-4 (**c**), and Gal-9 (**d**) with the presence of 2.00 μM to 200 mM cellobiose.

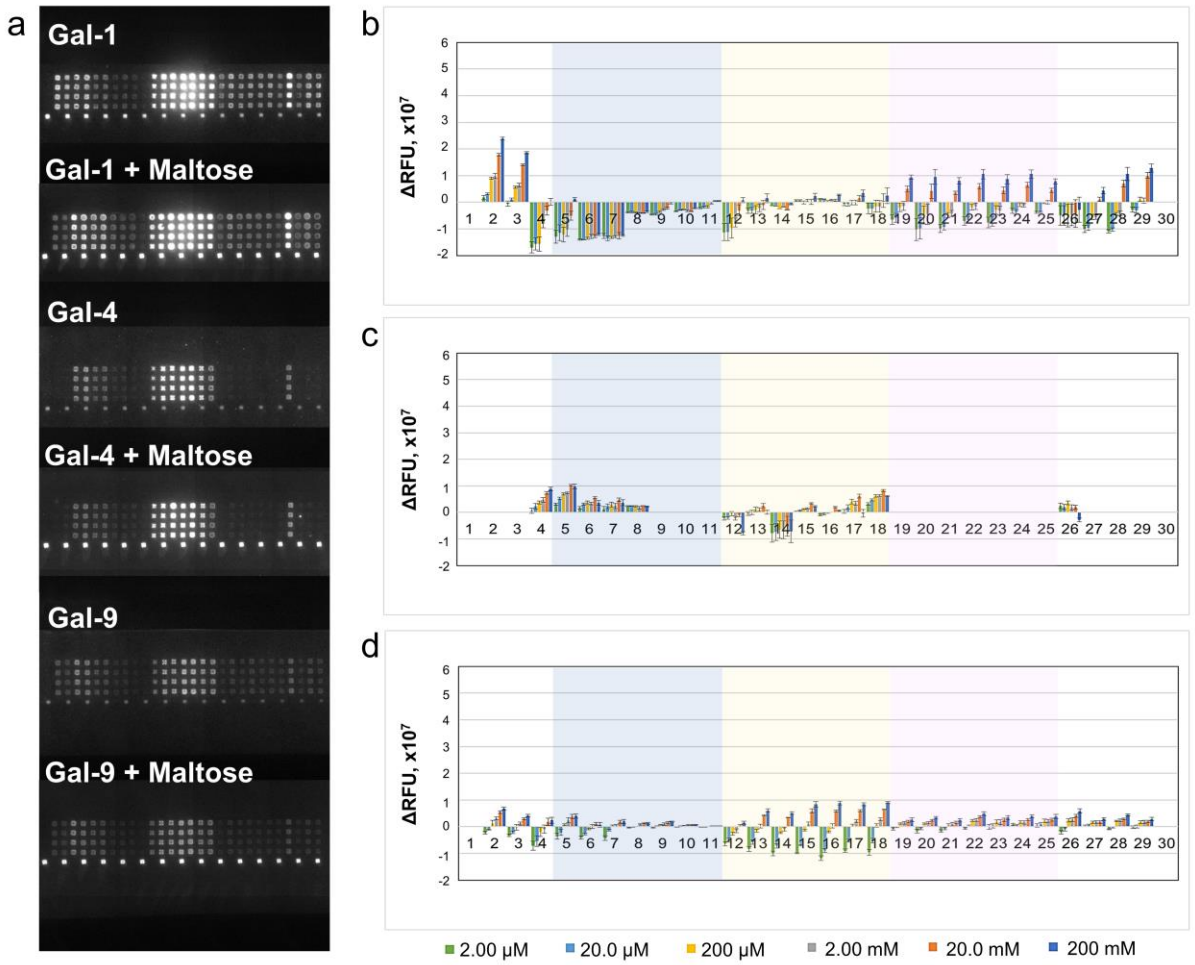

Figure S9. Fluorescence image of microarray chip is taken after treatment of 10.0 μg/mL lectin solution with 200 mM maltose (**a**) and relative binding properties of 200 μM core M1 α-DG glycopeptides Gal-1 (**b**), Gal-4 (**c**), and Gal-9 (**d**) with the presence of 2.00 μM to 200 mM maltose.

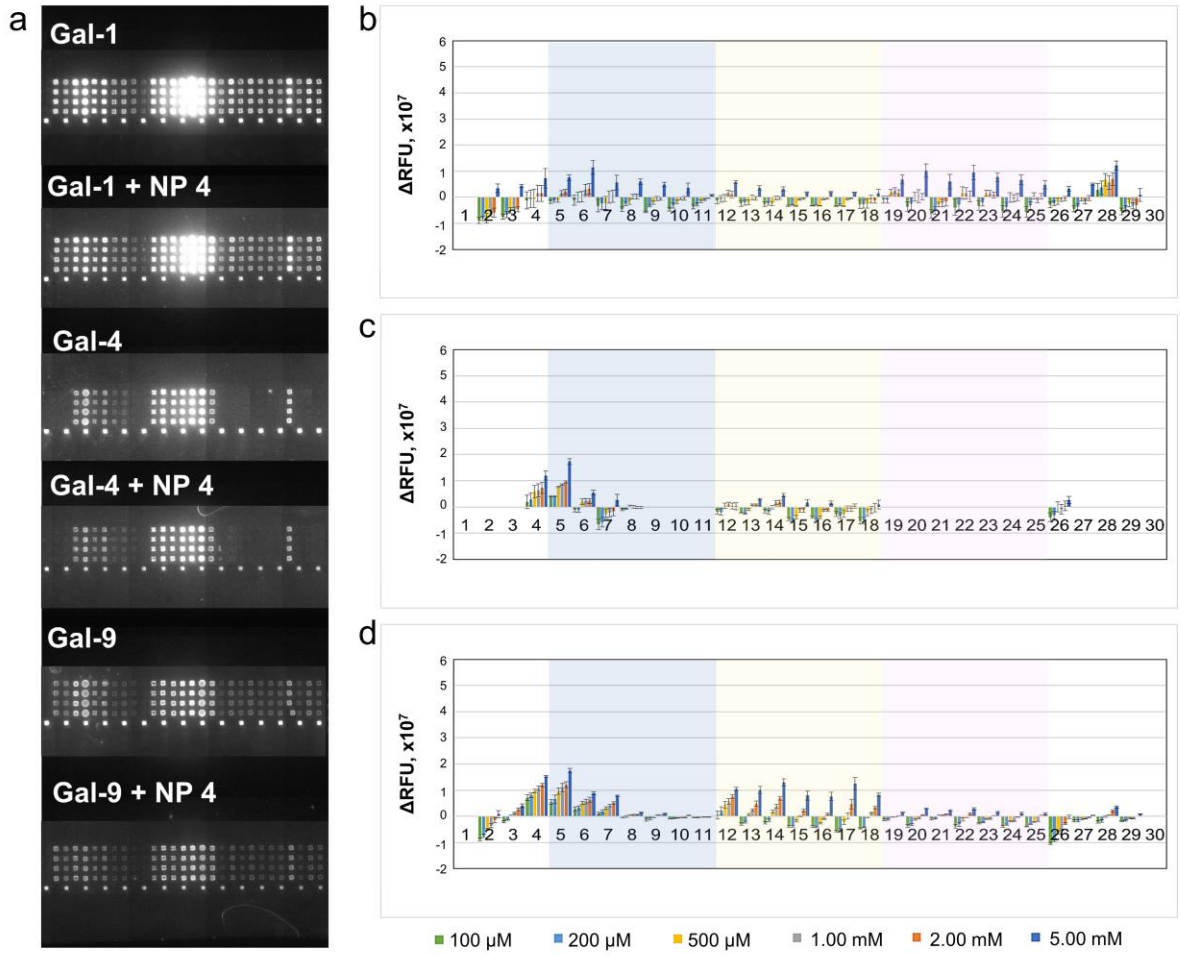

Figure S10. Fluorescence image of microarray chip is taken after treatment of 10 μg/mL lectin solution with 5.00 mM acetylated peptide **4** (**a**) and relative binding properties of 200 μM core M1 α-DG glycopeptides with Gal-1 (**b**), Gal-4 (**c**), and Gal-9 (**d**) with the presence of 100 μM to 5.00 mM acetylated peptide **4**.

2.5 References

1. T. Zhao, *et al.*, Hierarchy of Complex Glycomacromolecules: From Controlled Topologies to Biomedical Applications. *Biomacromolecules* **23**, 543–575 (2022).

2. T. Yue, B. B. Haab, Microarrays in Glycoproteomics Research. *Clin. Lab. Med.* **29**, 15–29 (2009).

3. T. Matsushita, *et al.*, A straightforward protocol for the preparation of high performance microarray displaying synthetic MUC1 glycopeptides. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 1105–1116 (2014).

4. A. Kuno, *et al.*, Evanescent-field fluorescence-assisted lectin microarray: a new strategy for glycan profiling. *Nat. Methods* **2**, 851–856 (2005).

5. G. Artigas, H. Hinou, F. Garcia-Martin, H.-J. Gabius, S.-I. Nishimura, Synthetic Mucin-Like Glycopeptides as Versatile Tools to Measure Effects of Glycan Structure/Density/Position on the Interaction with Adhesion/Growth-Regulatory Galectins in Arrays. *Chem. Asian J.* **12**, 159–167 (2017).

6. G. Artigas, *et al.*, Glycopeptides as Targets for Dendritic Cells: Exploring MUC1 Glycopeptides Binding Profile toward Macrophage Galactose-Type Lectin (MGL) Orthologs. *J. Med. Chem.* **60**, 9012–9021 (2017).

7. P. A. Guillen-Poza, *et al.*, Amplified Detection of Breast Cancer Autoantibodies Using MUC1-Based Tn Antigen Mimics. *J. Med. Chem.* **63**, 8524–8533 (2020).

8. H. Hinou, *et al.*, Synthetic glycopeptides reveal specific binding pattern and conformational change at O-mannosylated position of α-dystroglycan by POMGnT1 catalyzed GlcNAc modification. *Bioorg. Med. Chem.* **27**, 2822–2831 (2019).

9. T. Matsushita, H. Hinou, M. Kurogochi, H. Shimizu, S.-I. Nishimura, Rapid Microwave-Assisted Solid-Phase Glycopeptide Synthesis. *Org. Lett.* **7**, 877–880 (2005).

10. F. Garcia-Martin, H. Hinou, T. Matsushita, S. Hayakawa, S.-I. Nishimura, An efficient protocol for the solid-phase synthesis of glycopeptides under microwave irradiation. *Org. Biomol. Chem.* **10**, 1612–1617 (2012).

11. J. Iwaki, J. Hirabayashi, Carbohydrate-Binding Specificity of Human Galectins: An Overview by Frontal Affinity Chromatography. *Trends Glycosci. Glycotechnol.* **30**, SE137–SE153 (2018).

12. P. V Murphy, S. André, H.-J. Gabius, The Third Dimension of Reading the Sugar Code by Lectins: Design of Glycoclusters with Cyclic Scaffolds as Tools with the Aim to Define Correlations between Spatial Presentation and Activity. *Mol.* **18** (2013).

13. C. M. Guardia, *et al.*, Structural basis of redox-dependent modulation of galectin-1 dynamics and function. *Glycobiology* **24**, 428–441 (2014).

14. Y.-M. Go, D. P. Jones, Thiol/disulfide redox states in signaling and sensing. *Crit. Rev. Biochem. Mol. Biol.* **48**, 173–181 (2013).

15. S. R. Stowell, C. M. Arthur, R. D. Cummings, C. L. Feasley, Alkylation of galectin-1 with iodoacetamide and mass spectrometric mapping of the sites of incorporation. *Methods Mol. Biol.* **1207**, 51–62 (2015).

16. S. R. Stowell, *et al.*, Human galectin-1, -2, and -4 induce surface exposure of phosphatidylserine in activated human neutrophils but not in activated T cells. *Blood* **109**, 219–227 (2006).

17. H. Horie, T. Kadoya, Galectin-1 plays essential roles in adult mammalian nervous tissues. Roles of oxidized galectin-1. *Glycoconj. J.* **19**, 479–489 (2002).

18. T. Horlacher, *et al.*, Determination of carbohydrate-binding preferences of human galectins with carbohydrate microarrays. *Chembiochem* **11**, 1563–1573 (2010).

19. S. H. Stalnaker, *et al.*, Site Mapping and Characterization of O-Glycan Structures on α-Dystroglycan Isolated from Rabbit Skeletal Muscle*. *J. Biol. Chem.* **285**, 24882–24891 (2010).

20. M. I. Nielsen, *et al.*, Galectin binding to cells and glycoproteins with genetically modified glycosylation reveals galectin-glycan specificities in a natural context. *J. Biol. Chem.* **293**, 20249–20262 (2018).

21. M. C. Rodriguez, *et al.*, Thermodynamic Switch in Binding of Adhesion/Growth Regulatory Human Galectin-3 to Tumor-Associated TF Antigen (CD176) and MUC1 Glycopeptides. *Biochemistry* **54**, 4462–4474 (2015).

22. F. G. FitzGerald, *et al.*, TF-containing MUC1 glycopeptides fail to entice Galectin-1 recognition of tumor-associated Thomsen-Freidenreich (TF) antigen (CD176) in solution. *Glycoconj. J.* **37**, 657–666 (2020).

23. E. V Chandrasekaran, *et al.*, Novel interactions of complex carbohydrates with peanut (PNA), Ricinus communis (RCA-I), Sambucus nigra (SNA-I) and wheat germ (WGA) agglutinins as revealed by the binding specificities of these lectins towards mucin core-2 O-linked and N-linked glycans a. *Glycoconj. J.* **33**, 819–836 (2016).

24. J. Hirabayashi, *et al.*, Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. *Biochim. Biophys. Acta* **1572**, 232–254 (2002).

25. A. Leppänen, S. Stowell, O. Blixt, R. D. Cummings, Dimeric galectin-1 binds with high affinity to alpha2,3-sialylated and non-sialylated terminal N-acetyllactosamine units on surface-bound extended glycans. *J. Biol. Chem.* **280**, 5549–5562 (2005).

26. A. Borgert, B. L. Foley, D. Live, Contrasting the conformational effects of α-O-GalNAc and α-O-Man glycan protein modifications and their impact on the mucin-like region of alpha-dystroglycan. *Glycobiology* **31**, 649–661 (2021).

27. R. D. Cummings, J. D. Esko, Principles of Glycan Recognition (2009).

28. M. J. Moure, *et al.*, Selective 13C-Labels on Repeating Glycan Oligomers to Reveal Protein Binding Epitopes through NMR: Polylactosamine Binding to Galectins. *Angew. Chemie Int. Ed.* **60**, 18777–18782 (2021).

29. S. R. Stowell, *et al.*, Galectin-1, -2, and -3 exhibit differential recognition of sialylated glycans and blood group antigens. *J. Biol. Chem.* **283**, 10109–10123 (2008).

30. A. Tsouka, *et al.*, Probing Multivalent Carbohydrate-Protein Interactions With On-Chip Synthesized Glycopeptides Using Different Functionalized Surfaces. *Front. Chem.* **9** (2021).

31. L. Gauthier, B. Rossi, F. Roux, E. Termine, C. Schiff, Galectin-1 is a stromal cell ligand of the pre-B cell receptor (BCR) implicated in synapse formation between pre-B and stromal cells and in pre-BCR triggering. *Proc. Natl. Acad. Sci.* **99**, 13014 LP – 13019 (2002).

32. R. P. M. DINGS, *et al.*, beta-Sheet is the bioactive conformation of the anti-angiogenic anginex peptide. *Biochem. J.* **373**, 281–288 (2003).

33. M. Bozzi, *et al.*, Probing the stability of the “naked” mucin-like domain of human α-dystroglycan. *BMC Biochem.* **14**, 15 (2013).

34. S. Rangappa, *et al.*, Effects of the multiple O-glycosylation states on antibody recognition of the immunodominant motif in MUC1 extracellular tandem repeats. *Medchemcomm* **7**, 1102–1122 (2016).

35. F. Marcelo, *et al.*, Delineating binding modes of Gal/GalNAc and structural elements of the molecular recognition of tumor-associated mucin glycopeptides by the human macrophage galactose-type lectin. *Chemistry* **20**, 16147–16155 (2014).

36. T. K. Dam, T. A. Gerken, C. F. Brewer, Thermodynamics of Multivalent Carbohydrate−Lectin Cross-Linking Interactions: Importance of Entropy in the Bind and Jump Mechanism. *Biochemistry* **48**, 3822–3827 (2009).

37. P. von Hundelshausen, K. Wichapong, H.-J. Gabius, K. H. Mayo, The marriage of chemokines and galectins as functional heterodimers. *Cell. Mol. Life Sci.* **78**, 8073–8095 (2021).

38. J. Kopitz, S. Ballikaya, S. André, H.-J. Gabius, Ganglioside GM1/Galectin-Dependent Growth Regulation in Human Neuroblastoma Cells: Special Properties of Bivalent Galectin-4 and Significance of Linker Length for Ligand Selection. *Neurochem. Res.* **37**, 1267–1276 (2012).

39. A.-K. Ludwig, *et al.*, Design–functionality relationships for adhesion/growth-regulatory galectins. *Proc. Natl. Acad. Sci.* **116**, 2837 LP – 2842 (2019).

40. T. J. Kutzner, *et al.*, How altering the modular architecture affects aspects of lectin activity: case study on human galectin-1. *Glycobiology* **29**, 593–607 (2019).

Chapter 3.

Interaction between Gal-1 and α-DG core M1 glycopeptides by nuclear magnetic resonance

3.1 Introduction

3.1.1. Protein nuclear magnetic resonance (NMR) and Gal-1

Nuclear magnetic resonance (NMR) spectroscopy is a powerful analytical tool for determining protein structure, conformation, dynamics, and interactions from dilute protein-ligand solutions to living cells at the atomic level, similar to X-ray crystallography and cryogenic electron microscopy (1). Protein NMR structural and interaction studies usually involve (i) preparation of isotope-labeled protein or ligand, (ii) and data collection and analysis, mainly assignment of isotopic atoms (^{1}H, ^{13}C, ^{15}N, etc.) (2)(3). This technique can reveal a wide range of affinities (K_d in micro- to millimolar range) thus suitable for the characterization of carbohydrate-protein interactions (4).

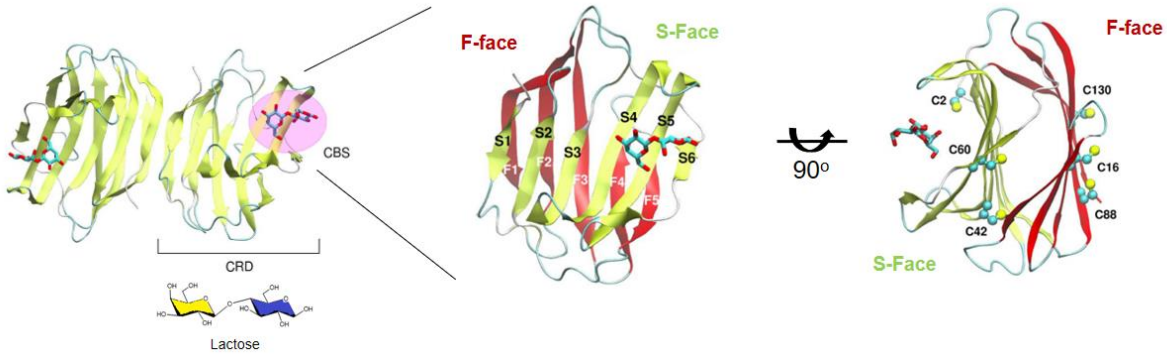

Figure 1. Interaction of lactose with Gal-1 (7).

Gal-1 (monomeric form, 14 kDa) is the first member of galectin family (5). The carbohydrate-recognition domain (CRD) of Gal-1 is consist of two anti-parallel β-sheets: S-Face (canonical carbohydrate-binding site: β1, β10, β3-6) and F-face (noncarbohydrate-binding site: β11, β2, β7-9). Gal-1 preferentially binds to β-galactose (Lactose (Lac) or N-acetyllactosamine (LacNAc)) structures on *N*- or *O*-linked glycans, Fig.1 (6). Among galectins, Gal-1 is unique because it contains several cysteine residues, thus sensitive to oxidation (Fig. 2) (7). Human Gal-1 dimerizes (homodimerization equilibrium, K_d = 2–7 μM) noncovalently through hydrophobic interactions between the monomers on its N- and C-terminal residues

which affects its biological activities by maintaining the reduced form of Gal-1 (Fig. 2) (8). Due to this, the multivalency of Gal-1 is suitable for facilitating cell-to-cell communication and triggering biosignals by forming “lattices” on the extracellular surface of different cell types (9)(10).

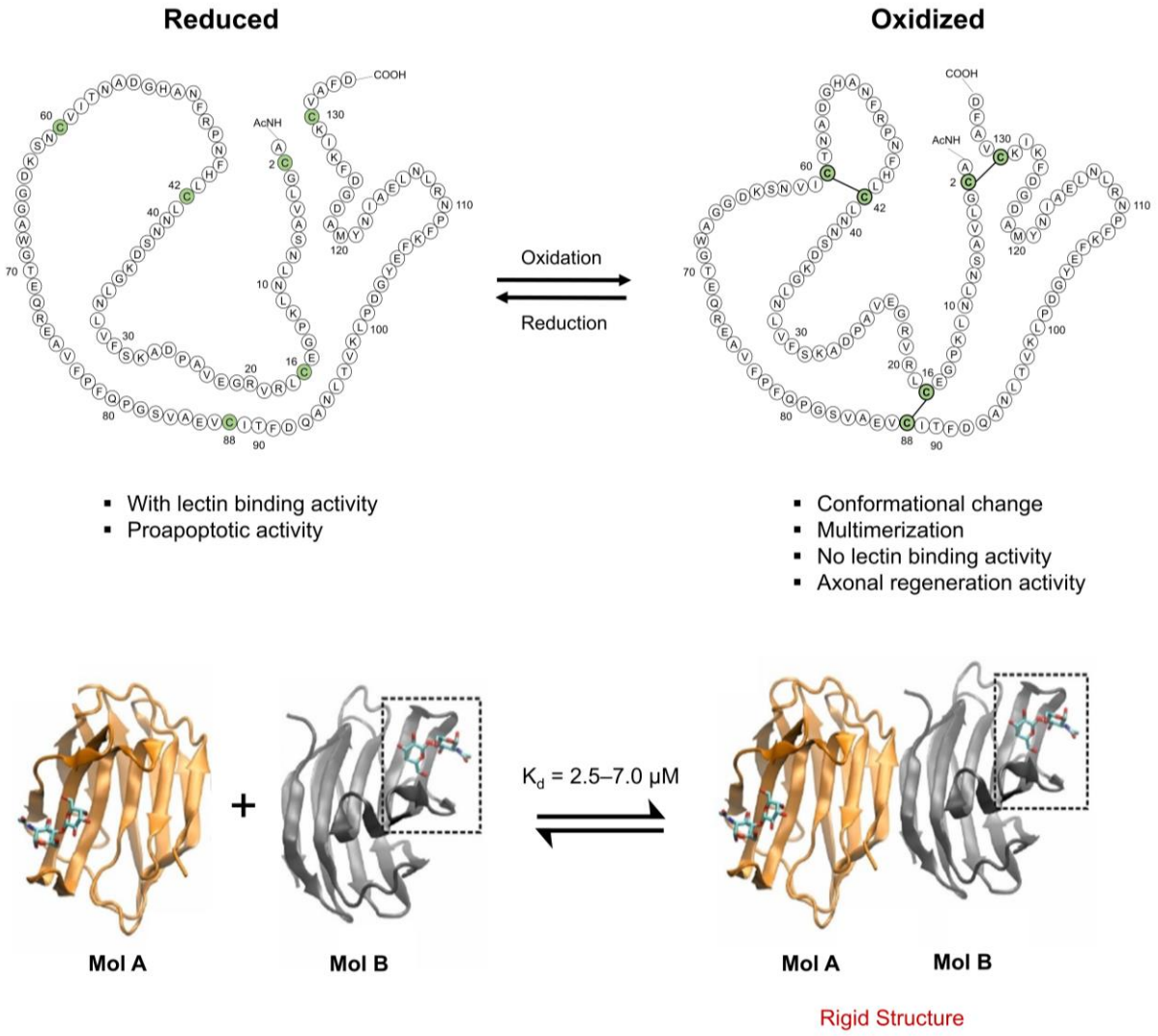

Figure 2. Amino acid sequence of Gal-1 and its dimerization (11)(12).

Both X-ray crystallography and NMR techniques revealed that Lac or LacNAc interaction occurs via hydrogen bonding (amino acid residues: H44, R48, H52, N61, and E71) and the stabilizing CH-π interactions (W68) on the S-face Fig. 3 (13)(14). Moreover, the conserved galectin-carbohydrate interaction, the galactose-terminated ligand is stabilized by CH-π interactions between the galactose ring and the aromatic ring of W68 (15).

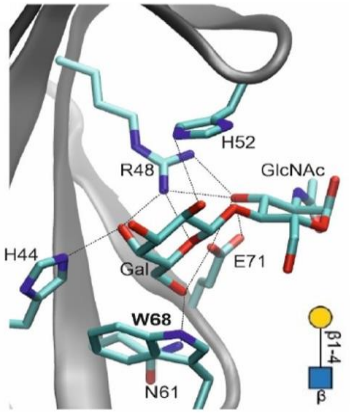

Figure 3. Crystal structure of recombinant human Gal-1 complexed with LacNAc (PDB: 1W6P) (12)(13).

In the previous chapters, our results revealed that galectins can interact with the core M1 glycopeptides of α-DG in two ways: with the LacNAc-terminated sidechains via its S-face and with the peptide backbone via specific hydrophobic interactions. To verify these hypothesized modes of Galectin binding with core M1 α-DG glycopeptides, we performed NMR experiments with Gal-1 and *N*-acetylated forms of glyco-amino acid **2**, peptide **4**, and glycopeptides **17**, and **24**.

3.2 Results and Discussion

The compounds **31**, **32**, **33**, and **34**, corresponding to **2**, **4**, **17**, and **24**, respectively, were synthesized Fig. 6a. For screening purposes, ^{15}N-labeled Gal-1 was prepared and tested under oxidizing and reducing conditions. Primarily, the Gal-1 ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) experiments were performed under oxidizing conditions (air oxidation or addition of oxidizing agents $CuSO_4$ or tetramethylazodicarboxamide) resulted in substantial aggregation of Gal-1 (16) with undetectable ^{1}H-^{15}N HSQC signals (SI Appendix, Fig. S1). In contrast, Gal-1 produced assignable signals under reducing conditions in the presence of DTT, as previously reported (17). The chemical shift perturbation of the 15N-labeled Gal-1 under

reducing conditions was analyzed with increasing concentration of methyl-Lac (positive control), **31**, **32**, **33**, and **34**.

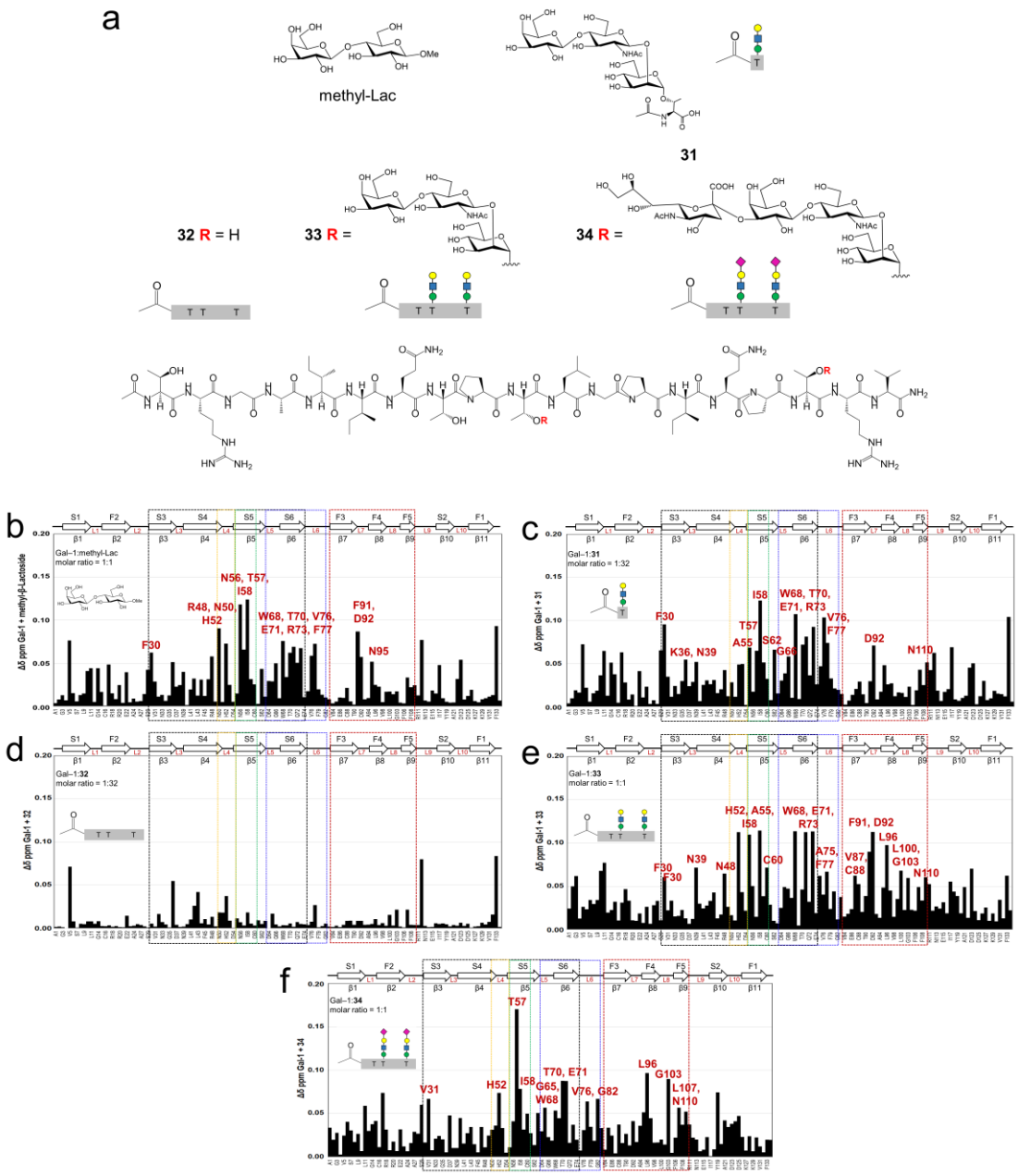

Figure 4. Structure of methyl-Lac, **31**, **32**, **33**, and **34** used for ^{1}H-^{15}N HSQC NMR studies (**a**). Chemical shift map (Δδ vs. sequence of Gal-1) is shown for methyl-Lac (**b**), glyco-amino acid **31** (**c**), unglycosylated α-DG glycopeptide **32** (**d**), LacNAc-terminated glycopeptide **33** (**e**), sialyl-LacNAc-terminated glycopeptide **34** (**f**). Chemical shift differences (Δδ) were calculated as $[(\Delta^{1}H)^{2}+(0.25\Delta^{15}N)^{2})]^{1/2}$. Solution condition were 20 mM potassium phosphate buffer, pH 6.9, 50 μM EDTA, with 10 mM DTT made up using a H_2O/D_2O (95:5%).

The presence of methyl-Lac shifted the 1H -15N HSQC signals corresponding to the residues on S3–L6 regions and induced a minor effect on the L7 – F4 regions of Gal-1, similar

to lactose in a concentration-dependent manner (Fig. 4b, SI Appendix Fig. S2a) (14). The LacNAc-terminated glyco-amino acid **31** induced limited perturbation in the chemical shift (Gal-1:**31** molar ratio of 1:1) (SI Appendix, Fig. S2b). When the molar ratio of Gal-1 to 31 was increased to 1:**32**, a remarkable perturbation of the chemical shift was observed at the S-face of Gal-1, similar to that of methyl-Lac (Fig. 4c). In contrast, unglycosylated peptide **32** failed to induce any significant deviation in the chemical shift at the tested concentrations (Fig. 4d, SI Appendix, Fig. S2c). Low concentration of LacNAc-terminated glycopeptide **33** (Gal-1:**33** molar ratio of 1:1) resulted in alterations of the chemical shifts at S-face (S4–S6 loop (L6)) and on the F-face (F3–F5) (Fig. 4e). Similar regions in Gal-1 were shifted by sialyl-LacNAc-terminated glycopeptide **34** (Gal-1:**34** molar ratio of 1:1); however, reduced deviations were observed on S4 and F3 regions (Fig. 4f). At a molar ratio greater than 1:1 for both **33** and **34** a significant reduction of Gal-1 signal intensities were observed (SI Appendix, Fig. S2d,e). This was attributed to Gal-1 oligomerization via cross-linking due to LacNAc binding to the S-face, and the specific hydrophobic binding formed between the peptide scaffold of α-DG and the F-face region of Gal-1. These supramolecular lectin-ligand complexes were undetectable by NMR, similar to glyco-amino acid **31** at high concentration (1:64) (SI Appendix, Fig. S2b). This phenomenon was also evident on Gal-1 binding with several multi-LacNAc-containing ligands such as galactorhamnogalacturonate, as previously demonstrated (18).

Comparing the perturbation pattern of the S-face of Gal-1 with methyl-Lac and glyco-amino acid **31**, the latter showed a relatively large deviation in the S6 region (W68, T70, E71, R73, V76, and F77) (Fig. 4b,c and 5a, SI Appendix, Fig. S3). The perturbation in the S6 region became more pronounced with the addition of LacNAc-terminated glycopeptide **33**, and the perturbation range was found to extend significantly to the F3–F4 regions (V87, C88, F91, D92, and L96) (Fig 4e and SI Appendix, Fig. S3). This mechanism is different from the

hydrophobic interaction of Gal-1 with the extracellular protein λ5 of the pre-BCR and CXCL4. When the pre-B cells are co-cultured with Gal-1+ stromal cells, the major element of the binding interaction of λ5 lies around the central α-helical structure (29RWGRFLLQRGS39), which makes hydrophobic contact with Gal-1 F-face (D102, Y104, E105, F106, L107, N110), triggering pre-BCR relocalization at the pre-B/stromal cell synapse, and concomitantly modifying the sugar ligand specificity and lattice interactions of Gal-1 (19)(20). Recently, the heterodimerization of galectins and chemokines via the glycan-independent interactions was found to occur, including Gal-1/CXCL4, which acts as an important modulator of inflammation (21). CXCL4 binding to Gal-1 induced changes in S6, F4, and F5 strands affecting the carbohydrate-binding region (S3 to S6 strands), which altered the Gal-1 affinity and specificity to its glycan ligands (22).

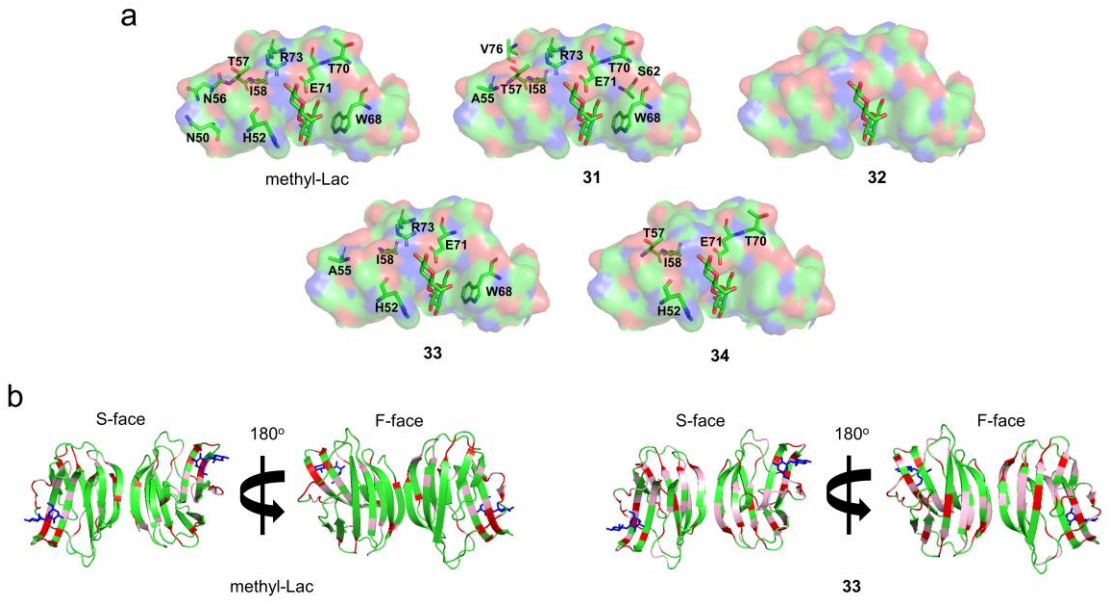

Figure 5. S-face focused electrostatic map of Gal-1 monomer (PDB access code: 1GZW) with the most perturbed amino acid residues by methyl-Lac, glyco-amino acid **31**, peptide **32**, LacNAc-terminated glycopeptide **33,** and sialyl-LacNAc-terminated glycopeptide **34** (**a**). The crystal structure of Gal-1 (PDB access code: 1GZW) is shown with the largest Δδ values highlighted in red (Δδ ≥ 0.05) and pink (0.025≥ Δδ <0.05) for methyl-Lac (**C**) and compound **33** (**b**). For orientation, lactose is presented in all models.

Interestingly, the perturbation with amino acid residues on the L4 region was also significantly enhanced in the presence of glycopeptide **33**. As a result, the aromatic side chains

of H52 and W68 recognized galactose, and the ionic couple of E71 and R73 interacted with 3-OH and N-acetyl groups of GlcNAc residue became more prominent (Fig. 4e and 5a; and SI Appendix, Fig. S3). This indicated that glycopeptide **33** could simultaneously interact with both the S- and F-face of Gal-1 via the S6 intermediate loop (Fig. 4e and 5a; and SI Appendix, Fig. S3). Contrastingly, **34** was predicted to lose its interaction with galactose and W68 because of the presence of sialic acid, as a result of pronounced deviation of the central T57 and adjacent I58 on the S5 region, which impacted the surrounding areas S4–L6 and F3–F4 loop (L8) regions (Fig. 4f and 5a; and SI Appendix, Fig. S3). The residue I58 is responsible for the intra-protein hydrophobic interaction between the S- and F-face and showed significant perturbation in all the experiments except for the unglycosylated peptide **32**.

3.3 Conclusion

Overall, our chemical shift perturbation mapping experiments agreed with our microarray results, where Gal-1 demonstrated high affinity to LacNAc-terminated core M1 glycopeptides via S- (protein-carbohydrate) and F-face (protein-peptide interaction), whose binding was decreased upon extension with α2,3-sialic acid. Although Gal-1 presented affinity to compounds **2** and **4** in a microarray study in the nanomolar range, the NMR experiments suggest that these ligands (**31** and **32**) exhibit binding above the millimolar range. This difference readily reflected the unique binding property of Gal-1 on a bulk surface such as a microarray, where ligands were immobilized at a high density for which *cis*-binding is likely to occur (23).

3.4 Experimental section

3.4.1 Materials

The pET-GST/TEV-LGVLFQGP-hLGALS1 vector was purchased from VectorBuilder, Japan. Other commercially available solvents and reagents were purchased from Sigma-Aldrich (USA), Tokyo Chemical Industry (Tokyo, Japan), Wako Pure Chemical Industries (Osaka, Japan), Kokusan Chemical (Tokyo, Japan), or Watanabe chemical (Osaka, Japan) and used without purification, unless stated.

3.4.2 Methods

Preparation of 15N-enriched gal-1

^{15}N Gal-1 was expressed in Escherichia coli cells BL21(DE3) from pET-GST/TEV-LGVLFQGP-hLGALS1 vector. The component cells were transformed with an expression vector utilizing the heat shock method. After one night of incubation on LB Agar plate (containing 100 μg/mL ampicillin), a colony harboring the expression construct was selected and then inoculated unto 5 mL Luria Broth (LB) with antibiotic overnight at 37ºC with shaking. The bacterial pellet was then resuspended to a fresh 500 mL M9-labeled (^{15}N-NH4Cl as nitrogen source) medium containing ampicillin. After that, culture was grown at 37ºC for at least 4 h until OD600 = 0.6 – 1.2. The cells were then induced with 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) overnight at 25ºC. Bacteria expressing 15N Gal-1 was lysed with a sonicator (for 5 min, 25% duty cycle. 5 output control) in buffer containing 50 mM PBS, pH 7.4, 150 mM NaCl, 2 mM DTT. 15N Gal-1 was purified using GSH Sepharose beads, cleaved with HRV3CC protease, dialysis, and gel filtration chromatography. Lectin purity as checked by 1D SDS-PAGE. Lectin-binding activity was analyzed by ^{15}N-HQSC with methyl-β-lactose. The oxidized form of Gal-1 was also prepared using the same protocol without the presence of DTT.

NMR Experiments.

The ^{15}N-labeled Gal-1 was prepared at 100 µM concentration in 20 mM potassium phosphate buffer at pH 6.9, 50 µM EDTA and 10 mM DTT, made up using H2O/D2O (95:5%). A 0.05 up to 64 equivalents of ligands were titrated with Gal-1. The 1H-15N Heteronuclear single quantum coherence (HSQC) NMR experiments were carried out at 25ºC on Bruker 600 AVANCE spectrophotometer equipped with QXI probes and z-axis pulse field gradient units. A gradient sensitivity-enhanced version of the two-dimensional 1H-15N HSQC experiment (128 scans per transient) was applied with 256 (t1) x 2048 (t2) complex data points in 15N and 1H dimensions, respectively. Raw data were converted using TopSpin 2.1 and were analyzed using NMRFAM-Sparky (24).

3.4.3. Supplementary information

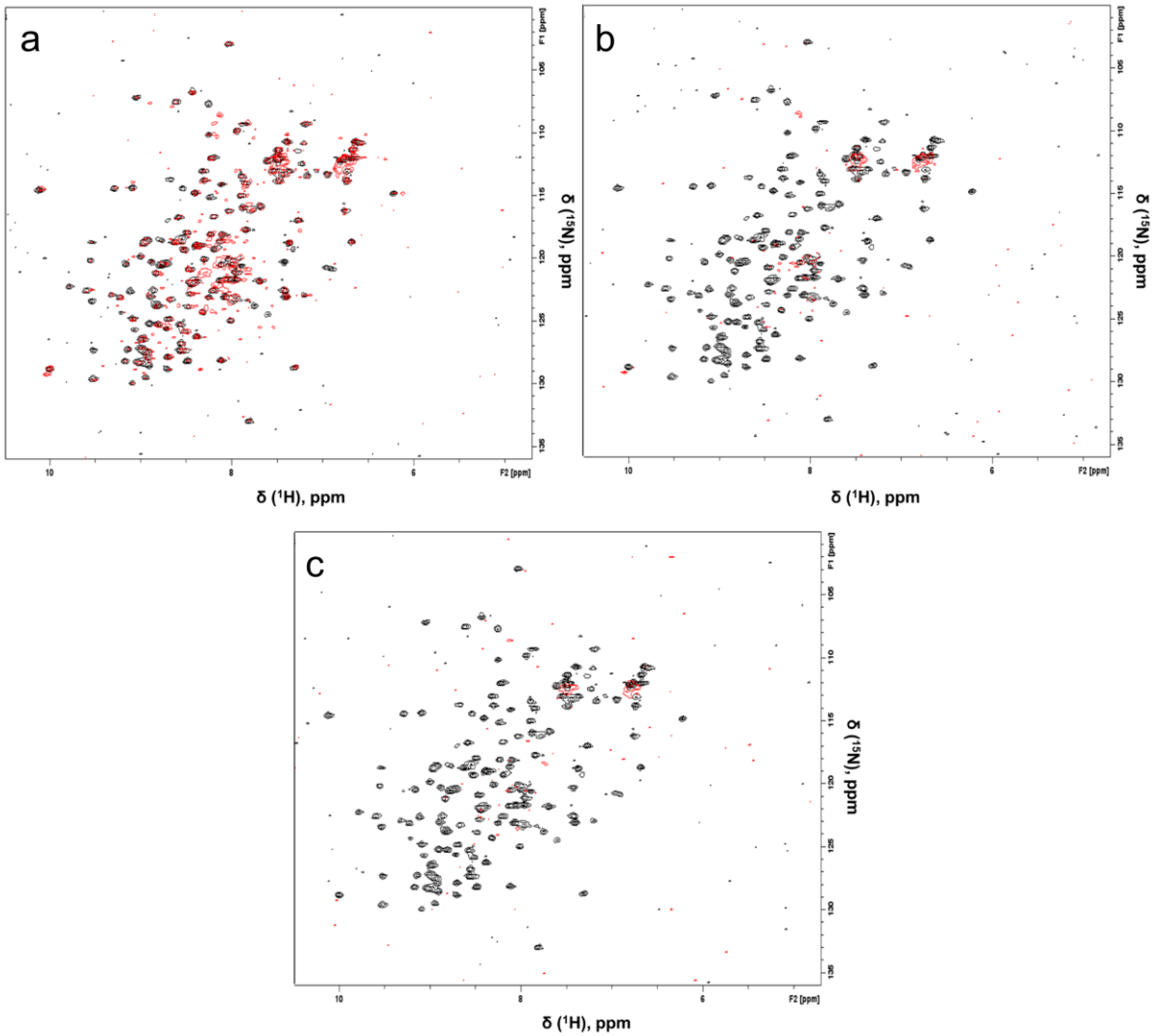

Figure S1. Overlaid ^{1}H-^{15}N HSQC spectra of 100 µM Gal-$1_{reduced}$ (black) with 100 µM Gal-$1_{oxidized}$ (red) using air (**a**), 6.9 µM $CuSO_4$ (**b**), and 1.0 mM tetramethylazodicarboxamide (**c**).

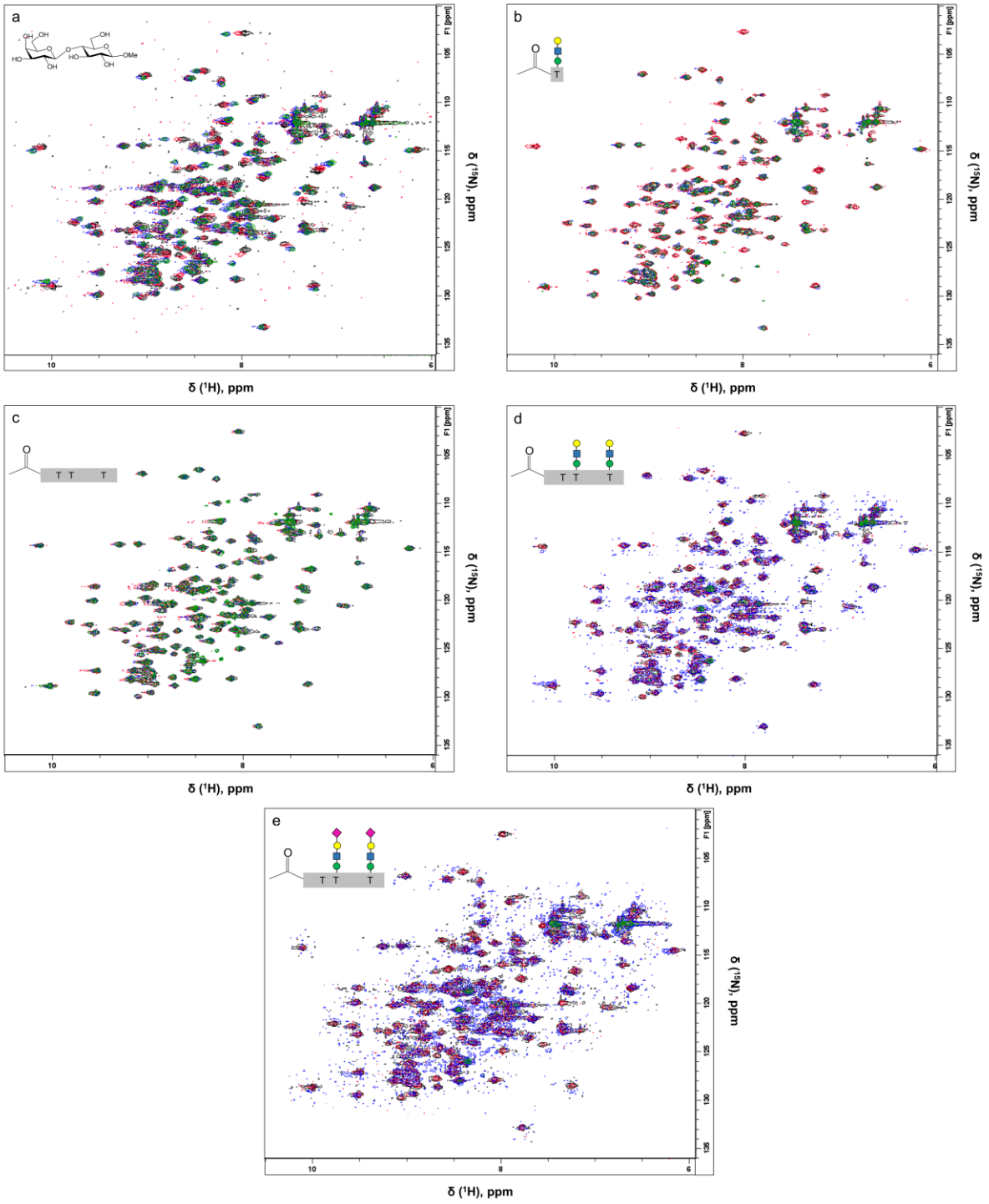

Figure S2. Overlaid ^{1}H-^{15}N HSQC spectra of reduced Gal-1 (100 µM) alone (**black** peaks) and in the presence of methyl-β-lactoside (molar ratio 1:1 **blue** peaks, 1:16 **red** peaks, and 1:64 **green** peaks, **a**), LacNAc-terminated glycoamino acid **31** (molar ratio 1:1 **blue** peaks, 1:32 **red** peaks, and 1:64 **green** peaks **b**), unglycosylated peptide **32** (molar ratio 1:1 **blue** peaks, 1:16 **red** peaks, and 1:64 **green** peaks **c**), LacNAc-terminated glycopeptide acid **33** (molar ratio 1:0.125 **blue** peaks, 1:1 **red** peaks, and 1:2 **green** peaks **d**), sialyl-LacNAc-terminated glycopeptide **34** (molar ratio 1:0.125 **blue** peaks, 1:1 **red** peaks, and 1:2 **green** peaks **e**), Solution condition were 20 mM potassium phosphate buffer, pH 6.9, 50 µM EDTA, with 10 mM DTT made up using a H_2O/D_2O (95:5%).

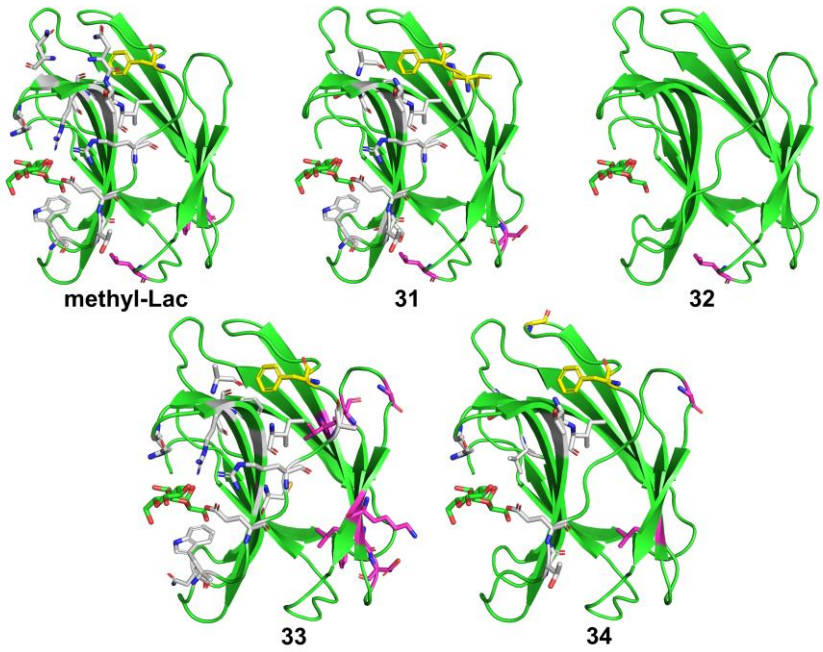

Figure S3. Close-up view of Gal-1 monomer S6 region binding site of methyl-β-lactose, glycoamino acid **31**, nonglycosylated peptide **32**, glycopeptides **33** and **34**. Peptide backbone (green); most perturb amino acid residues in S-Face (gray); most perturb amino acid in F-face (violet). For orientation, lactose (C and O atoms are highlighted in green and red, respectively) is presented.

3.5 References

1. Y. Hu, *et al.*, NMR-Based Methods for Protein Analysis. *Anal. Chem.* **93**, 1866–1879 (2021).

2. M. R. O'Connell, R. Gamsjaeger, J. P. Mackay, The structural analysis of protein–protein interactions by NMR spectroscopy. *Proteomics* **9**, 5224–5232 (2009).

3. O. Cala, F. Guillière, I. Krimm, NMR-based analysis of protein–ligand interactions. *Anal. Bioanal. Chem.* **406**, 943–956 (2014).

4. C. A. Bewley, S. Shahzad-ul-Hussan, Characterizing carbohydrate-protein interactions by nuclear magnetic resonance spectroscopy. *Biopolymers* **99**, 796–806 (2013).

5. S. H. Barondes, *et al.*, Galectins: a family of animal beta-galactoside-binding lectins. *Cell* **76**, 597–598 (1994).

6. I. Camby, M. Le Mercier, F. Lefranc, R. Kiss, Galectin-1: a small protein with major functions. *Glycobiology* **16**, 137R-157R (2006).

7. C. P. Modenutti, J. I. B. Capurro, S. Di Lella, M. A. Martí, The Structural Biology of Galectin-Ligand Recognition: Current Advances in Modeling Tools, Protein Engineering, and Inhibitor Design . *Front. Chem.* **7**, 823 (2019).

8. J. M. Romero, M. Trujillo, D. A. Estrin, G. A. Rabinovich, S. Di Lella, Impact of human galectin-1 binding to saccharide ligands on dimer dissociation kinetics and structure. *Glycobiology* **26**, 1317–1327 (2016).

9. C. F. Brewer, M. C. Miceli, L. G. Baum, Clusters, bundles, arrays and lattices: novel mechanisms for lectin-saccharide-mediated cellular interactions. *Curr. Opin. Struct. Biol.* **12**, 616–623 (2002).

10. G. A. Rabinovich, M. A. Toscano, S. S. Jackson, G. R. Vasta, Functions of cell surface galectin-glycoprotein lattices. *Curr. Opin. Struct. Biol.* **17**, 513–520 (2007).

11. H. Horie, T. Kadoya, Galectin-1 plays essential roles in adult mammalian nervous tissues. Roles of oxidized galectin-1. *Glycoconj. J.* **19**, 479–489 (2002).

12. F. Tobola, *et al.*, Engineering the Ligand Specificity of the Human Galectin-1 by Incorporation of Tryptophan Analogues. *ChemBioChem* **23**, e202100593 (2022).

13. M. F. López-Lucendo, *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.* **343**, 957–970 (2004).

14. I. V Nesmelova, *et al.*, Lactose Binding to Galectin-1 Modulates Structural Dynamics, Increases Conformational Entropy, and Occurs with Apparent Negative Cooperativity. *J. Mol. Biol.* **397**, 1209–1230 (2010).

15. C. Meynier, F. Guerlesquin, P. Roche, Computational studies of human galectin-1: role of conserved tryptophan residue in stacking interaction with carbohydrate ligands. *J. Biomol. Struct. Dyn.* **27**, 49–58 (2009).

16. C. M. Guardia, *et al.*, Structural basis of redox-dependent modulation of galectin-1 dynamics and function. *Glycobiology* **24**, 428–441 (2014).

17. I. V Nesmelova, M. Pang, L. G. Baum, K. H. Mayo, 1H, 13C, and 15N backbone and side-chain chemical shift assignments for the 29 kDa human galectin-1 protein dimer. *Biomol. NMR Assign.* **2**, 203–205 (2008).

18. H. Ippel, *et al.*, Intra- and intermolecular interactions of human galectin-3: assessment by full-assignment-based NMR. *Glycobiology* **26**, 888–903 (2016).

19. L. Elantak, *et al.*, Structural basis for galectin-1-dependent pre-B cell receptor (pre-BCR) activation. *J. Biol. Chem.* **287**, 44703–44713 (2012).

20. J. Bonzi, *et al.*, Pre-B cell receptor binding to galectin-1 modifies galectin-1/carbohydrate affinity to modulate specific galectin-1/glycan lattice interactions. *Nat. Commun.* **6**, 6194 (2015).

21. V. Eckardt, *et al.*, Chemokines and galectins form heterodimers to modulate inflammation. *EMBO Rep.* **21**, e47852 (2020).

22. L. Sanjurjo, *et al.*, Chemokines modulate glycan binding and the immunoregulatory activity of galectins. *Commun. Biol.* **4**, 1415 (2021).

23. J. Pai, *et al.*, Carbohydrate microarrays for screening functional glycans. *Chem. Sci.* **7**, 2084–2093 (2016).

24. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

Chapter 4.

Gal-1 trans-bridge α-DG core M1 glycopeptides and laminins in microarray

4.1. Introduction

4.1.1. Laminin, muscular dystrophy, and Gal-1

Laminins (~400–900kDa) are one of the major glycoprotein components of the basal lamina that anchor cell membrane receptors, such as integrins, syndecan, and extended core M3 structures of dystroglycan (DG). It provide cellular stability and play a pivotal role in establishing the normal tissue architecture by holding cells and tissues together (1)(2)(3), Fig. 1.

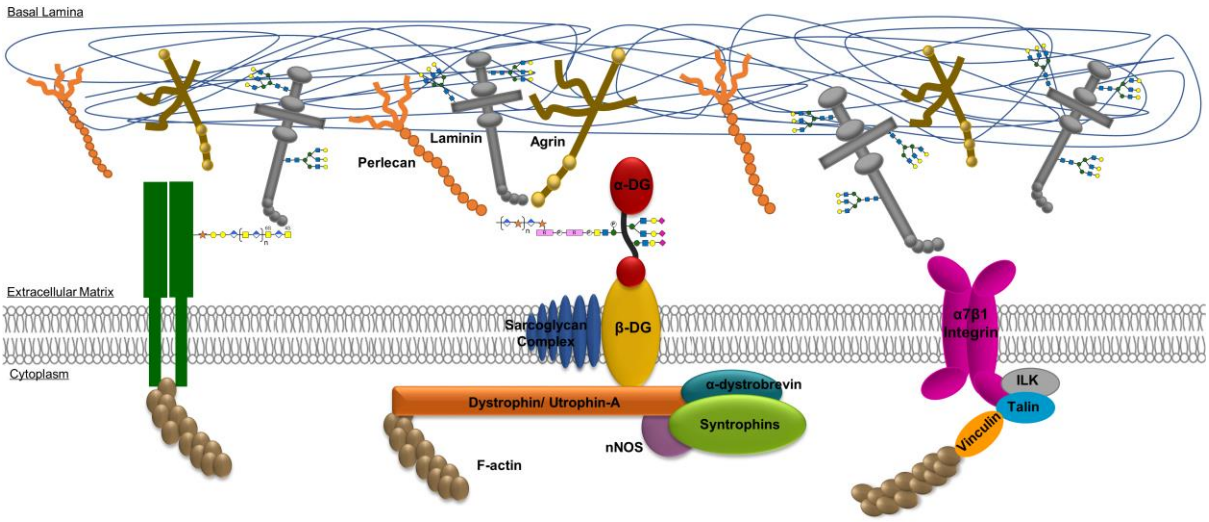

Figure 1. Cell membrane receptors of skeletal muscle to the extracellular matrix of basal lamina.

Structurally, laminins are made up of α, β, and γ chains connected by three disulfide-linked polypeptides (4). The composition of each chain is used to designate the isoform of this structurally diverse family, Fig. 2 (5). Aside from maintaining the architecture and stability of basement membranes, laminins also separate and connect different tissues, provide adjacent cells with mechanical scaffold, and trigger biosignals by interacting with growth factors which are important for development and tissue homeostasis (6)(7). In adult skeletal muscle, α7β1 integrin plays a pivotal role in assisting myoblast adhesion to laminin-111, -211 and -221 during repair, while laminin-121 and -511 exhibited interaction with α7β1 and α6β1,

respectively, supporting neurite outgrowth (8)(9)(10). Deficiency of some laminin isoform can lead to congenital muscular dystrophy (11).

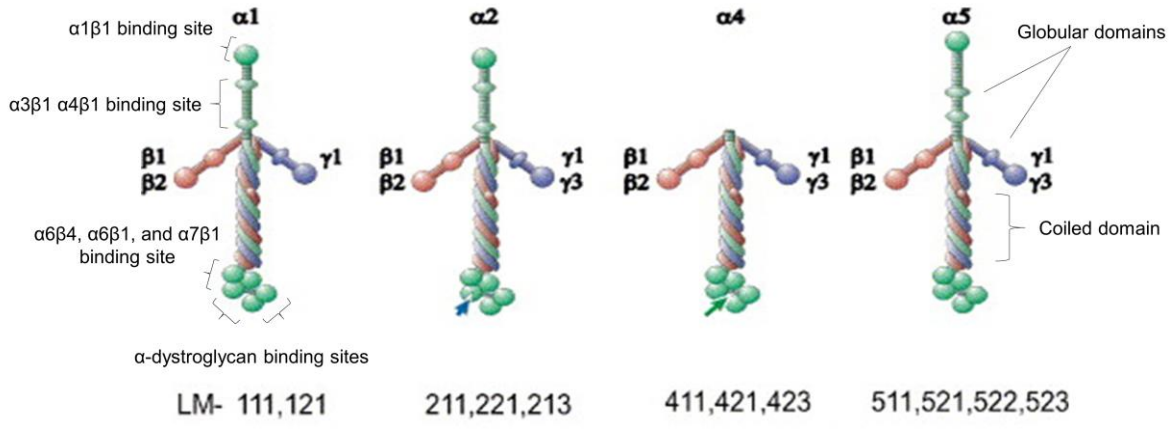

Figure 2. Representative human laminins. Numbers below the laminins indicate the trimer composition. Greek letters with numbers indicate chains. The blue arrow indicates cleavage that results in a fragment that remains associated non-covalently with the trimer. The green arrow indicates cleavges found to generate free fragments. (5)

α-DG is one of the well-established ligand of laminins and other LG-containing extracellular matrix proteins with high binding affinities (K_d up to nM range) (12)(13). Notably, hypoglycosylation of α-DG induces severe disruption of DG–laminin binding in patients with Fukuyama congenital muscular dystrophy (FCMD) and muscle-eye brain (MEB) disease (14)(15).

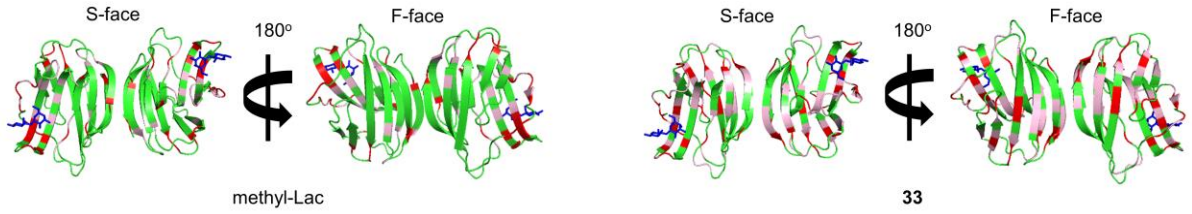

Figure 3. The crystal structure of Gal-1 (PDB access code: 1GZW) is shown with the largest Δδ values highlighted in red (Δδ ≥ 0.05) and pink (0.025≥ Δδ <0.05) for methyl-Lac (**C**) and compound **33** (**b**) based on HSQC results. For orientation, lactose is presented in all models.

Preclinical studies revealed that exogenous application of Gal-1 is an exciting therapeutic potential for treating certain forms of muscular dystrophy(16); however, the precise mechanisms underpinning this phenomenon remain unclear (17). In the basal lamina, the polyLacNAc units of laminins (111, 211, and 511) serve as ligands for Gal-1 ($K_d \approx 10^{-6}$M), which maintains its carbohydrate-binding activity in the extracellular matrix (ECM) (17)(18)(19)(20). During muscle repair, Gal-1 binds directly to laminin and α7β1 integrin modulating the late myoblast fusion (21). In addition, Gal-1 can dissociate from laminin and trans-bridge olfactory neurons, forming neuronal aggregates in vitro (22). In the previous chapter, we demonstrated that Gal-1 can interact with via the S- and F-face of its CRD, Fig. 3. Thus, we investigated whether non-labeled Gal-1 could *trans*-bridge various fluorescently labeled laminins and core M1 glycopeptides of α-DG using microarray technique.

4.2. Results and Discussion

As previously reported, laminin -111 and -211 could not interact with the α-DG core M1 glycopeptides at pH 7.4 (23). Here, Gal-1 revealed trans-bridging capabilities, linking laminin-111, -121, -211, and -221 (but little -511) and core M1 α-DG glycopeptides (Fig. 3 and SI Appendix, Fig. S1). Gal-1 exhibited low laminin-bridging capabilities with glyco-amino acids **1**–**3**. Interestingly, the unglycosylated α-DG peptide **4** and core M1 α-DG glycopeptides (**5**–**11**), but not MUC1 peptide **30**, could anchor laminin in the presence of Gal-1. The position (**5** vs. **6** vs. **7**) and density (**5**–**7** vs. **8**–**10** vs. **11**) of the core M1 structure along the peptide scaffold affected the *trans*-bridging activity (Fig. 4; and SI Appendix, Fig. S1). This can be attributed to the hydrophobic interaction of Gal-1 with peptide **4 – 11**, which was not dependent on the S-face of the CRD, as discussed previously. In Gal-1 homodimer, F-face and the canonical S-face sites are oriented on the same side, respectively (Fig. 3). This

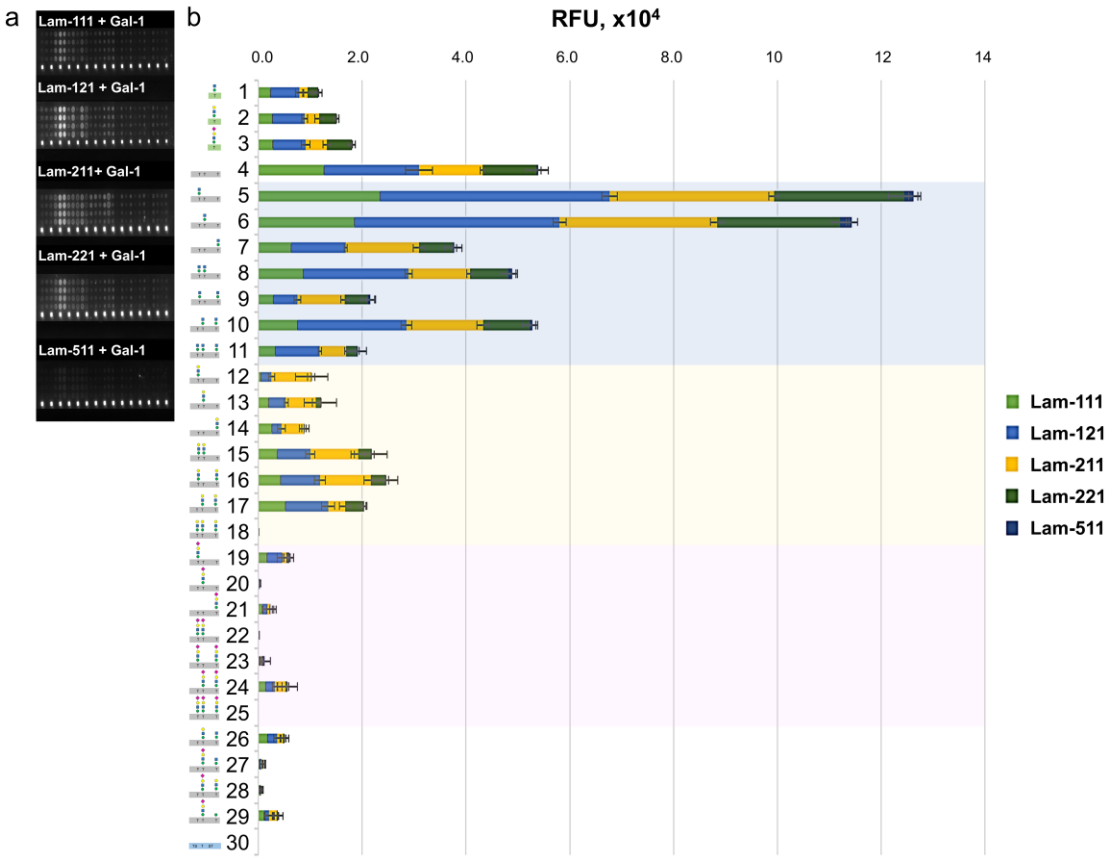

Figure 4. Fluorescence image of microarray chip (**a**) and stacked chart of signal intensities (**b**) of 200 μM core M1 α-DG glycoconjugates taken after treatment with 32.0 μg/mL Cy3-laminin–unlabeled-Gal-1 solution for 30 min. The relative interaction of 200 μM α-DG unglycosylated peptides with 0.10 μg/mL to 32.0 μg/mL Cy3-laminin–unlabeled-Gal-1 solution is in Supplementary Fig. S1.

structural feature of the Gal-1 homodimer is expected to allow simultaneous interaction with *N*-glycans on the laminin at the S-face and with α-DG peptide at the F-face. The remarkable signal intensities demonstrated by compounds **5** and **6**, which did not show significant interaction with Gal-1, suggest that the *N*-glycan ligand interaction of laminin at the Gal-1 S-face induced a structural and property change in F-face, as observed in NMR study. This result is consistent with the enhanced interaction of Gal-1 with compounds **5** and **6** upon addition of methyl-Lac (Chapter 3, Fig. 6). The previously reported conformational change in the peptide backbone associated with the addition of core M1 could also contribute to this activity change (23). This unique trans-bridging activity of Gal-1 was also observed during the pre-B/stromal cell synapse formation, where Gal-1 binds to pre-B cell λ5 (via protein-protein interaction) and

integrins (via protein-carbohydrate interaction), driving the clustering and activation of pre-B cell receptors (BCRs) (24).

A decrease in the *trans*-bridging activity of Gal-1 was observed upon galactosylation of the core M1 structures (**12**–**18**, **28**). In this case, the two S-face of Gal-1 homodimer preferentially bind to poly-LacNAc units of laminins (*cis*-binding) rather than one S-face to LacNAc unit of laminin and one S-face to LacNAc-terminated core M1 α-DG glycopeptide (*trans*-bridging). Further extension of the LacNAc core M1 structures with sialic acid (**19**–**25**, **27**–**29**) completely abolished the *trans*-bridging activity of Gal-1 with laminin.

Among the evaluated laminins, laminin-511 was not trans-bridged by Gal-1 to the α-DG peptide and core M1 glycopeptides. The α-chain of laminins possesses a unique structure and is the most glycosylated compared to β- and γ-chains (25)(26). This result indicates that different profile or levels of glycans is present in different α-chain of laminin isoforms which is necessary for the trans-bridging activity of Gal-1 with the prepared ligands. However, the participation of other poly-LacNAc structures located in the β- and γ-chains cannot be neglected.

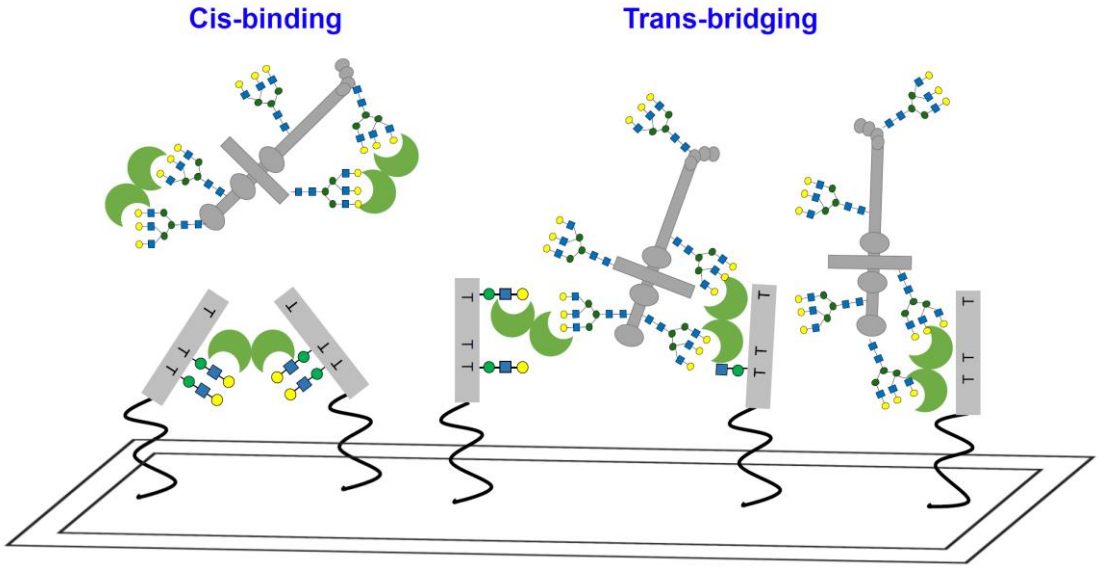

Figure 5. The *cis*-binding and *trans*-bridging activities of Gal-1 with α-DG core M1 glycoconjugates and laminin *in situ*.

4.3. Conclusion

Gal-1 effectively cross-linked various laminins (111, 121, 211, and 221, but not 511) to the α-DG peptide and core M1 glycopeptides in the array. This Gal-1 trans-bridging activity was completely abrogated by the α2,3-sia extension of the LacNAc core M1 glycopeptides, indicating that Gal-1 binding to the polyLacNAc side chain of laminin is preferred over α-DG core M1 sialyl-LacNAc ligands (Fig. 5). These results suggest a novel insight into the potential mechanism by which Gal-1 prevents progression of muscle during muscular dystrophy, specifically FCMD where hypoglycaosylation of dystroglycan core M3 structure is highly observed.

4.4. Experimental section

4.4.1. Materials

Laminins (111, 121, 211, 221, and 511) and Cy3 NHS antibody labeling kit were procured from Biolamina and BroadPharm, respectively.

4.4.2. Methods

Labeling of laminins

Cy3-NHS was dissolved with 50 µL of anhydrous DMSO. An equal amount of laminin solution was added with a final concentration of 100 µg/mL. The mixture was incubated for 1 hr at room temperature with gentle mixing protected from light. The excess dye was removed and the buffer was changed utilizing the supplied desalting column. The Cy3-label laminins where then stored at 4^{o}C until use.

Laminin–Gal-1 Binding Assay

A 12 µL of premixed Cy3-laminin and unlabeled-Gal-1 with final concentration of 0.10 µg/mL was added to the printed test compounds. After 1 hr of incubation with the laminin-galectin solution at rt in a humidified chamber, the slide was rinsed with washing buffer and fluorescence intensity was measured with GlycoStation System. To determine whether additional interactions will be detected at higher laminin-Gal-1 concentrations, the reaction buffer solution was carefully removed and replaced with the next test concentration, incubated for 5 min at rt, washed with reaction buffer and then fluorescence intensity was obtained. This step was repeated until all the chosen test concentrations were completed (0.32-, 1.00-, 3.20-, 10.0-, and 32.0 µg/mL). For weak interaction, incubation with 32.0 µg/mL of laminin-Gal-1 solution was extended for 30 min. Slide images were captured and analyzed as described above.

Images of slides were captured in the presence of reaction buffer. The fluorescence intensities obtained were analyzed using ArrayVision software. For each spot, background

correction was applied to get the net intensity, and the average relative fluorescence unit (RFU) was plotted using Microsoft Excel software. The lectin specificity was identified from high and low RFU values, the error bars being the standard deviation.

4.4.3. Supplementary information

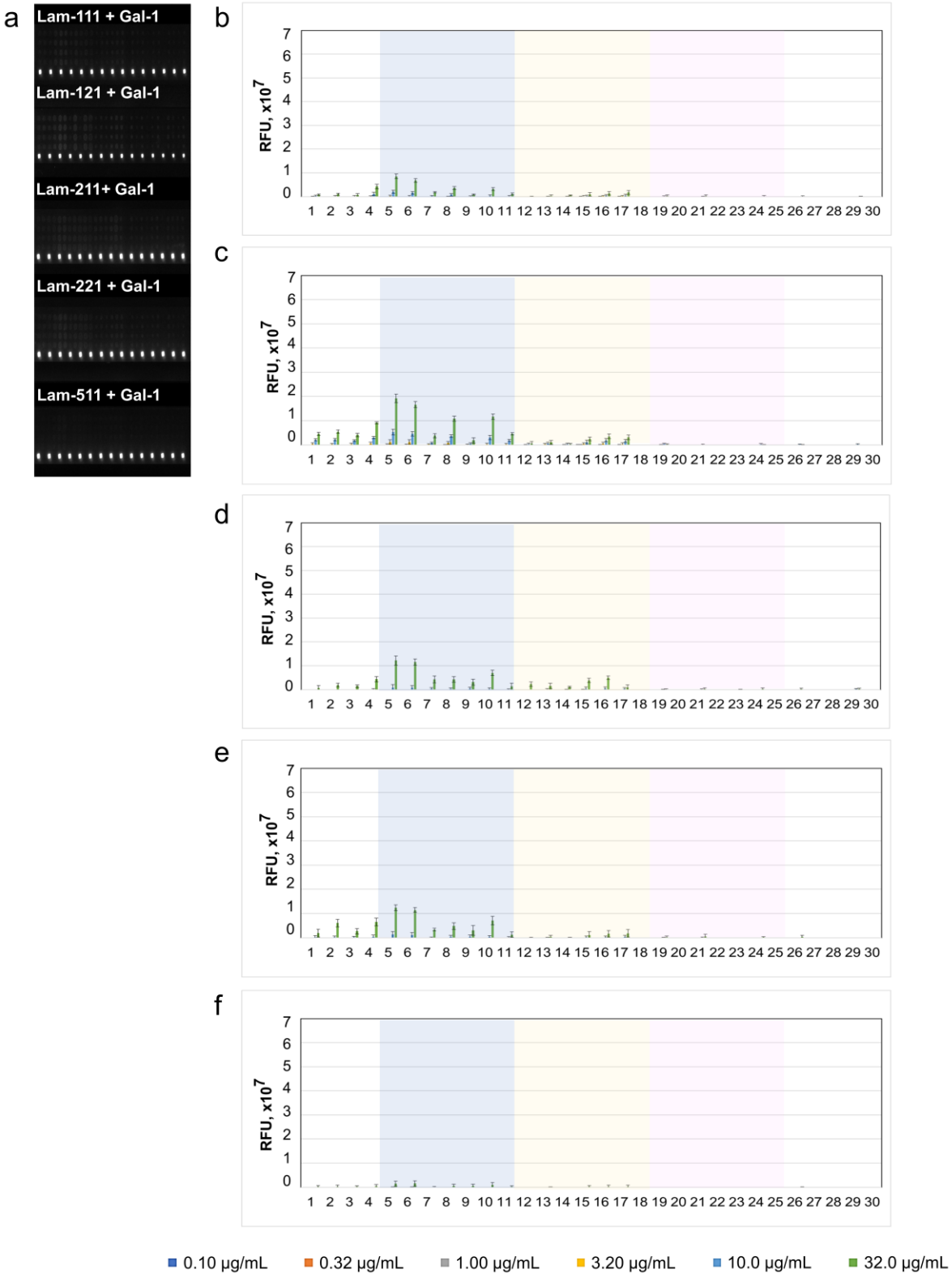

Figure S1. Fluorescence image of microarray chip is taken after treatment of 200 µM core M1 α-DG glycopeptides with 10.0 µg/mL laminin–Gal-1 solution (**a**) and relative binding properties of 200 µM core M1 of αDG with 0.10 to 32.0 µg/mL Laminin-111 and Gal-1 (**b**), Laminin-121 and Gal-1 (**c**), Laminin-211 and Gal-1 (**d**), Laminin-221 and Gal-1 (**e**), Laminin-511 and Gal-1 (**f**).

4.5. References

1. T. Yoshida-Moriguchi, K. P. Campbell, Matriglycan: a novel polysaccharide that links dystroglycan to the basement membrane. *Glycobiology* **25**, 702–713 (2015).

2. H. Renzhi, *et al.*, Basal lamina strengthens cell membrane integrity via the laminin G domain-binding motif of α-dystroglycan. *Proc. Natl. Acad. Sci.* **106**, 12573–12579 (2009).

3. M. Aumailley, N. Smyth, The role of laminins in basement membrane function. *J. Anat.* **193 (Pt 1**, 1–21 (1998).

4. M. Aumailley, The laminin family. *Cell Adh. Migr.* **7**, 48–55 (2013).

5. M. Aumailley, *et al.*, A simplified laminin nomenclature. *Matrix Biol.* **24**, 326–332 (2005).

6. K. Garg, M. D. Boppart, Influence of exercise and aging on extracellular matrix composition in the skeletal muscle stem cell niche. *J. Appl. Physiol.* **121**, 1053–1058 (2016).

7. J. Ishihara, *et al.*, Laminin heparin-binding peptides bind to several growth factors and enhance diabetic wound healing. *Nat. Commun.* **9**, 2163 (2018).

8. S. Crawley, *et al.*, The α7β1 Integrin Mediates Adhesion and Migration of Skeletal Myoblasts on Laminin. *Exp. Cell Res.* **235**, 274–286 (1997).

9. S. Plantman, *et al.*, Integrin-laminin interactions controlling neurite outgrowth from adult DRG neurons in vitro. *Mol. Cell. Neurosci.* **39**, 50–62 (2008).

10. T. Sasaki, *et al.*, Laminin-121—Recombinant expression and interactions with integrins. *Matrix Biol.* **29**, 484–493 (2010).

11. H. Kölbel, *et al.*, Identification of Candidate Protein Markers in Skeletal Muscle of Laminin-211-Deficient CMD Type 1A-Patients . *Front. Neurol.* **10** (2019).

12. S. H. Gee, *et al.*, Laminin-binding protein 120 from brain is closely related to the dystrophin-associated glycoprotein, dystroglycan, and binds with high affinity to the major heparin binding domain of laminin. *J. Biol. Chem.* **268**, 14972–14980 (1993).

13. L. R. Sheckler, L. Henry, S. Sugita, T. C. Südhof, G. Rudenko, Crystal structure of the second LNS/LG domain from neurexin 1alpha: Ca2+ binding and the effects of alternative splicing. *J. Biol. Chem.* **281**, 22896–22905 (2006).

14. D. E. Michele, *et al.*, Post-translational disruption of dystroglycan–ligand interactions in congenital muscular dystrophies. *Nature* **418**, 417–421 (2002).

15. M. Kanagawa, *et al.*, Residual laminin-binding activity and enhanced dystroglycan glycosylation by LARGE in novel model mice to dystroglycanopathy. *Hum. Mol. Genet.* **18**, 621–631 (2009).

16. P. M. Van Ry, R. D. Wuebbles, M. Key, D. J. Burkin, Galectin-1 Protein Therapy Prevents Pathology and Improves Muscle Function in the mdx Mouse Model of Duchenne Muscular Dystrophy. *Mol. Ther.* **23**, 1285–1297 (2015).

17. P. M. Van Ry, *et al.*, ECM-Related Myopathies and Muscular Dystrophies: Pros and Cons of Protein Therapies. *Compr. Physiol.* **7**, 1519–1536 (2017).

18. E. A. Pall, K. M. Bolton, J. M. Ervasti, Differential heparin inhibition of skeletal muscle alpha-dystroglycan binding to laminins. *J. Biol. Chem.* **271**, 3817–3821 (1996).

19. H. Yu, J. F. Talts, Beta1 integrin and alpha-dystroglycan binding sites are localized to

different laminin-G-domain-like (LG) modules within the laminin alpha5 chain G domain. *Biochem. J.* **371**, 289–299 (2003).

20. M. T. Elola, M. E. Chiesa, A. F. Alberti, J. Mordoh, N. E. Fink, Galectin-1 receptors in different cell types. *J. Biomed. Sci.* **12**, 13–29 (2005).

21. M. Gu, W. Wang, W. K. Song, D. N. Cooper, S. J. Kaufman, Selective modulation of the interaction of alpha 7 beta 1 integrin with fibronectin and laminin by L-14 lectin during skeletal muscle differentiation. *J. Cell Sci.* **107 (Pt 1**, 175–181 (1994).

22. N. K. Mahanthappa, D. N. Cooper, S. H. Barondes, G. A. Schwarting, Rat olfactory neurons can utilize the endogenous lectin, L-14, in a novel adhesion mechanism. *Development* **120**, 1373–1384 (1994).

23. H. Hinou, *et al.*, Synthetic glycopeptides reveal specific binding pattern and conformational change at O-mannosylated position of α-dystroglycan by POMGnT1 catalyzed GlcNAc modification. *Bioorg. Med. Chem.* **27**, 2822–2831 (2019).

24. B. Rossi, M. Espeli, C. Schiff, L. Gauthier, Clustering of Pre-B Cell Integrins Induces Galectin-1-Dependent Pre-B Cell Receptor Relocalization and Activation. *J. Immunol.* **177**, 796 LP – 803 (2006).

25. P. D. Yurchenco, *et al.*, The α chain of laminin-1 is independently secreted and drives secretion of its β- and γ-chain partners. *Proc. Natl. Acad. Sci.* **94**, 10189–10194 (1997).

26. D. N. W. Cooper, 荒田洋一郎, ガレクチン-1の分泌及びガレクチン-1とラミニンによる細胞間相互作用の調節. *Trends Glycosci. Glycotechnol.* **9**, 57–67 (1997).

Chapter 5.

Altering the modular architecture of galectin affects its binding with α-DG core M1 glycoconjugates

5.1. Introduction

5.1.1. Galectin protein engineering

The functional pairing of tissue lectins and cellular glycoconjugates is receiving increasing attention as these interactions are involved in a wide range of (patho)physiological processes (1–4). The glycan-encoded biochemical signals ("sugar code") are translated into their distinct cellular effects ("special meaning") by lectins ("the readers") (5). As driven by evolution, the diversity and specificity of this molecular recognition entail that structural and topological aspects on both sides influence this functional pairing. Lectins are protein receptors that consist of a carbohydrate recognition domain (CRD) to accommodate glycan epitopes with strong avidity. This capability efficiently facilitates cell-cell or cell-matrix adhesion (trans-bridging) or signal transduction (cis-crosslinking) which is critically dependent on lectin CRD's molecular presentation. Lectins are classified according to the protein fold by their CRD, which is diversified on the sequence and modular design level, with phylogenetic variation in each group, displaying distinct characteristics (6, 7), Fig.1.

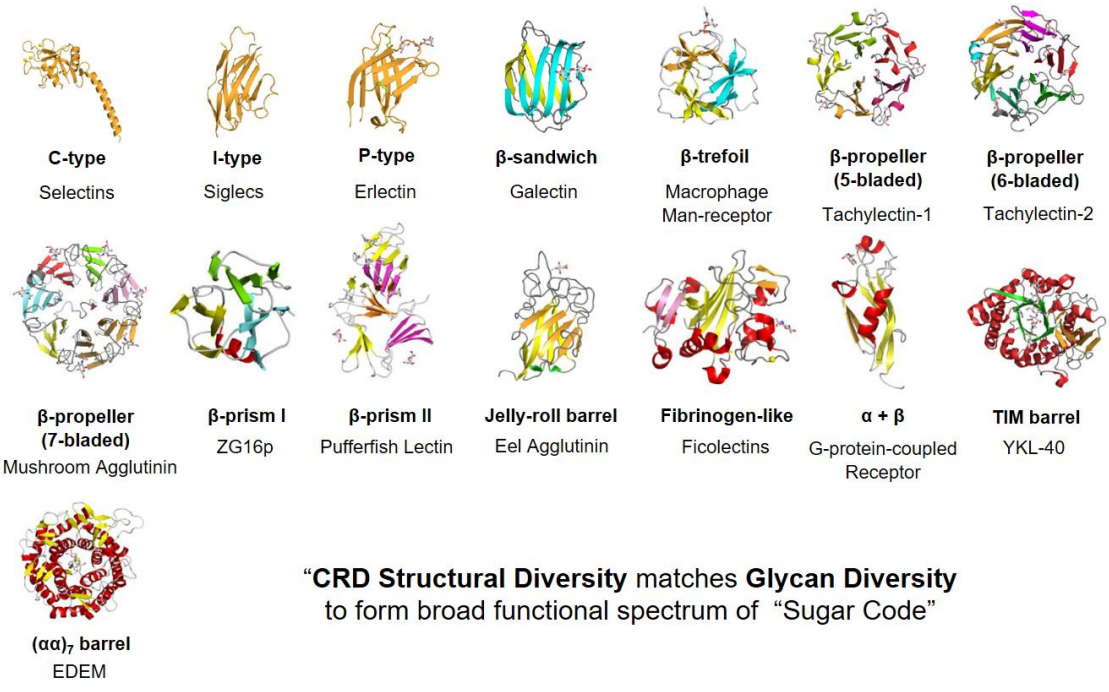

Figure 1. Diversity of folding within the superfamily of lectins (6)(7).

Focusing on adhesion/growth-regulatory galectins (Gal), this lectin family preferentially binds to galactose-terminated glycoconjugates. Naturally, this lectin family is present on three phylogenetically conserved structural designs (Fig. 2A): i) a non-covalent linked homodimer (prototype), ii) two nonidentical but homologous CRD associated by a peptide linker (tandem-repeat type), and iii) a CRD connected to an N-terminal tail with non-triple-helical collagen-like repeats (chimera-type). These natural forms of galectins can be further modified via alternative splicing and proteolytic cleavage, giving rise to multiple isoforms within this lectin family (8, 9). With this panel of galectin architecture, the CRD-counterreceptor interactions can form lattices of unique spatial characteristics, dependent on CRD presentation, eliciting post-binding events (10). Galectin CRDs are extensively studied; however, it is still unclear why galectin CRDs are limited to these types of modular designs. Thus, rational protein engineering is a promising method to delineate this natural preference and the impact of CRD structure alteration on lectin activity. Newly engineered galectins to probe the effects of CRD organization and spacing on receptor cross-linking were successfully tested using cells, glycoclusters, glycodendrimers, glycopolymers, and glycopeptides (11–14). This approach will further provide insights into the design-functionality correlation of the target specificity of wild-type galectins and the prospect of translational therapeutic applications of these human lectin variants.

For instance, several modifications of Gal-1, have been attempted to improve the oxidation sensitivity and bioactivity of the wild-type protein. A cysteine-less mutant of Gal-1 (C3/17/43/61/89/131S, CSGal-1), having similar functional and structural properties to the parent protein, was prepared as a stable substitute for wild-type Gal-1 (Gal-1WT) (15). Glycine-glycine (GG) linked Gal-1, displays Gal-1 as single chain bivalent galectin, demonstrated a 10-fold increase in apoptotic activity in both murine thymocytes and mature T cells the Gal-1WT (16). The Gal-1 tandem-repeat type mutant with 14 amino acid Gal-9 short

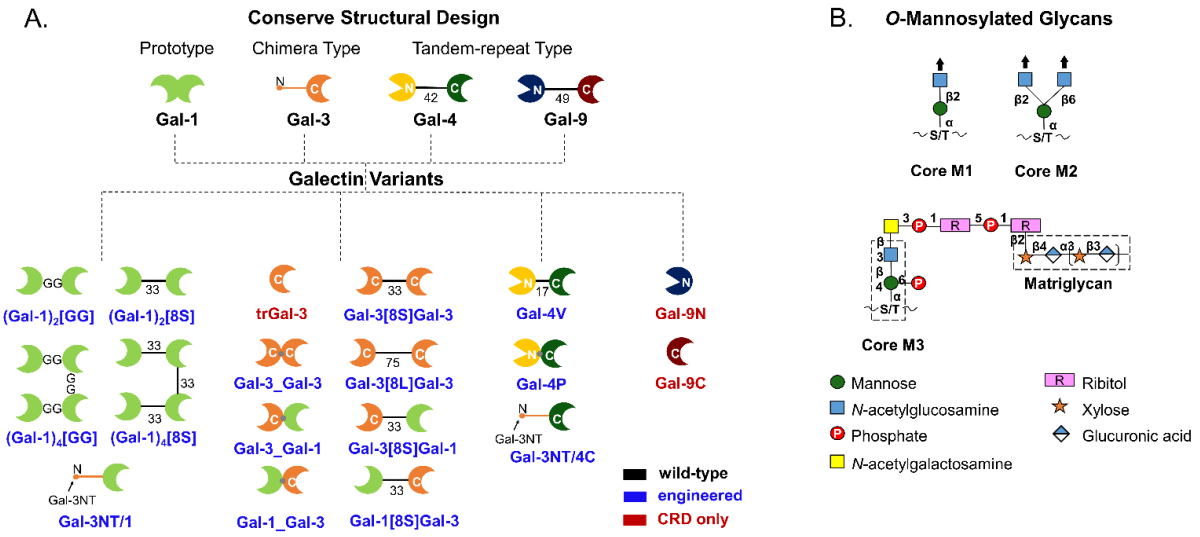

Figure 2. (A) Structural design of galectin wild-type and variants. [8S] = Gal-8, 33-aa peptide linker; [8L] = Gal-8, 74-aa peptide linker; Gal-3NT = Gal-3 N-terminal Tail; V = shorten Gal-4 linker; P = Prototype; N = N-terminal CRD; C= C-temrinal CRD. (B) Classification of *O*-Mannosylated glycans in α-DG.

isoform random-coil linker (Gal-1-9-1) exhibited 30-fold enhancement in triggering Jurkat E6-1 cell death over Gal-1 wild-type (17). In oysters and bay scallops, a Gal-1 isoform with quadruple-CRD was identified and is involved in the recognition, phagocytosis, and elimination of several pathogenic microbes in these invertebrates (18, 19). Incorporating this architectural design into human Gal-1 (homo-oligomer variants ((Gal-1)$_4$-GG/8S)) resulted in 25-fold higher cell-bridging activity than Gal-1WT as revealed by hemagglutination and erythrocyte aggregation experiments (12). Furthermore, a heterotetramer galectin variant having two Gal-1 and Gal-3 domains (G1/G3 Zipper) demonstrated higher apoptosis activity than Gal-1 and Gal-3 wild-type alone or in combination and thus can be a good therapeutic candidate in regulating innate and adaptive immunity (20, 21).

The dystroglycan (DG) is a transmembrane protein that provides a link between the extracellular matrix and intracellular cytoskeleton via the α-DG and β-DG, respectively (22). The α-DG is highly *O*-Mannosylated (*O*-Man) which plays an important role for its ability to bind basement membrane protein containing laminin G (LamG) domains (Fig. 2B) (23). The matriglycan of α-DG has been identified as the receptor for laminins, maintaining skeletal

muscle integrity and normal development of central nervous system (24). The disruption of α-DG glycosylation leads to various forms of muscular dystrophy which have been known for several decades, but there are still no effective treatments for this group of diseases (25). In our previous study, we demonstrated that LacNAc-terminated *O*-Man core M1 glycopeptide of α-DG fragment ^{372}TRGAIIQ**T**P**T**LGPIQP**T**RV^{390} interact with galectin that proceed via *cis*-binding (peptide- and carbohydrate-protein interactions) *in situ* (26). The systematic variation of headgroup, density, and position presentation on the peptide scaffold, and the type of galectin native structure affected the lectin affinity. The *trans*-bridging capability of wild-type Gal-1 with core M1 glycoconjugates and laminins was also demonstrated *in situ*. Here, we combined the α-DG glycopeptide and galectin variant libraries to analyze the impact of switching galectin design and valency on the binding process utilizing glycan microarray technology. Furthermore, we tested the impact of the alteration of Gal-1 structure (presence of linker, increasing the CRD per protein molecule and transformation to heterodimer) to the observed α-DG-laminin *trans*-bridging activities of the wild-type protein.

5.2. Results and Discussion

5.2.1. Galectin variant toolbox

Our galectin test panel encompasses the three types of galectin architectural design found in humans. We select prototype Gal-1, chimera-type Gal-3, and two tandem-repeated types Gal-4 and -9 as test groups (Fig. 1A). To examine the architecture-dependent functionality of galectins, engineering by modular transportation allowed us to switch galectin classes and access new combinations (4, 27–30). For Gal-1, an engineered covalently-linked homodimer, i.e., **(Gal-1)$_2$**, and homotetramer, i.e., **(Gal-1)$_4$**, of Gal-1 connected by dipeptide Gly-Gly (**GG**) and 33-aa linker of tandem repeat-type Gal-8 (**8S**), and a chimera variant

inspired obtained by transplanting the mono Gal-1 CRD to Gal-3's NT, i.e., **Gal-3NT/1** were prepared.

The wild-type Gal-3 was turned into a homodimer by directly linking C- and N-terminal amino acids of two Gal-3 CRDs (**Gal-3_Gal-3**) or by inserting tandem-repeat type presentation Gal-8 linker, i.e., 33-aa (8S) or 74-aa (8L), yielding **Gal-3[8S/L]Gal-3** variants. Heterodimers consisting with Gal-3 and Gal-1 CRDs were constructed. Here, the structural variables were the spatial order of the CRD, i.e., **Gal-1/-3** or **Gal-3/-1** (from the N to C terminus), in the absence or presence of a linker (**8S**).

For tandem-repeat type Gal-4, to address the issue of the relevance of the linker length for its binding capacity, two variants, that is with reduced size (**Gal-4V**) and complete linker removal mimicking a dimer of the proto-type group (**Gal-4P**) were designed and produced. In addition, Gal-4 C-terminal CRD was incorporated into a Gal-3-like design, i.e., **Gal-3NT/4**. While for Gal-9, CRDs with no linker, **Gal-9N/C** were constructed.

The Lac/LacNAc binding properties of these prepared galectin variants were found to be similar to their wild-type counterparts, thus proving the whole activity of all lectins (29–32). With this highly diverse galectin toolbox, an extensive comparative analysis of the three architectural designs of human galectins, including evaluating the impact of linker length and mode of CRD presentation, is achievable. These proteins were isolated under activity-preserving conditions and employed for growth regulation, bridging, and aggregation assays using cells and glycodendrimersomes (28, 33). Here we test the interaction profiles of this lectin family to an array-based screening with core M1 α-DG glycopeptides with distinct structural variations.

5.2.2. α-DG *O*-Man core M1 glycopeptide library

Surface immobilized glycopeptides in microarray serve as one of the effective tools to investigate galectin-carbohydrate interactions. The advantage of this method is one can determine the precise binding epitope of galectins and provide insight into the effect of other glycan moieties, position, and density along the peptide scaffold. Here we utilized our previously reported microarray platform, based on the 19 amino acid along the mucin-like

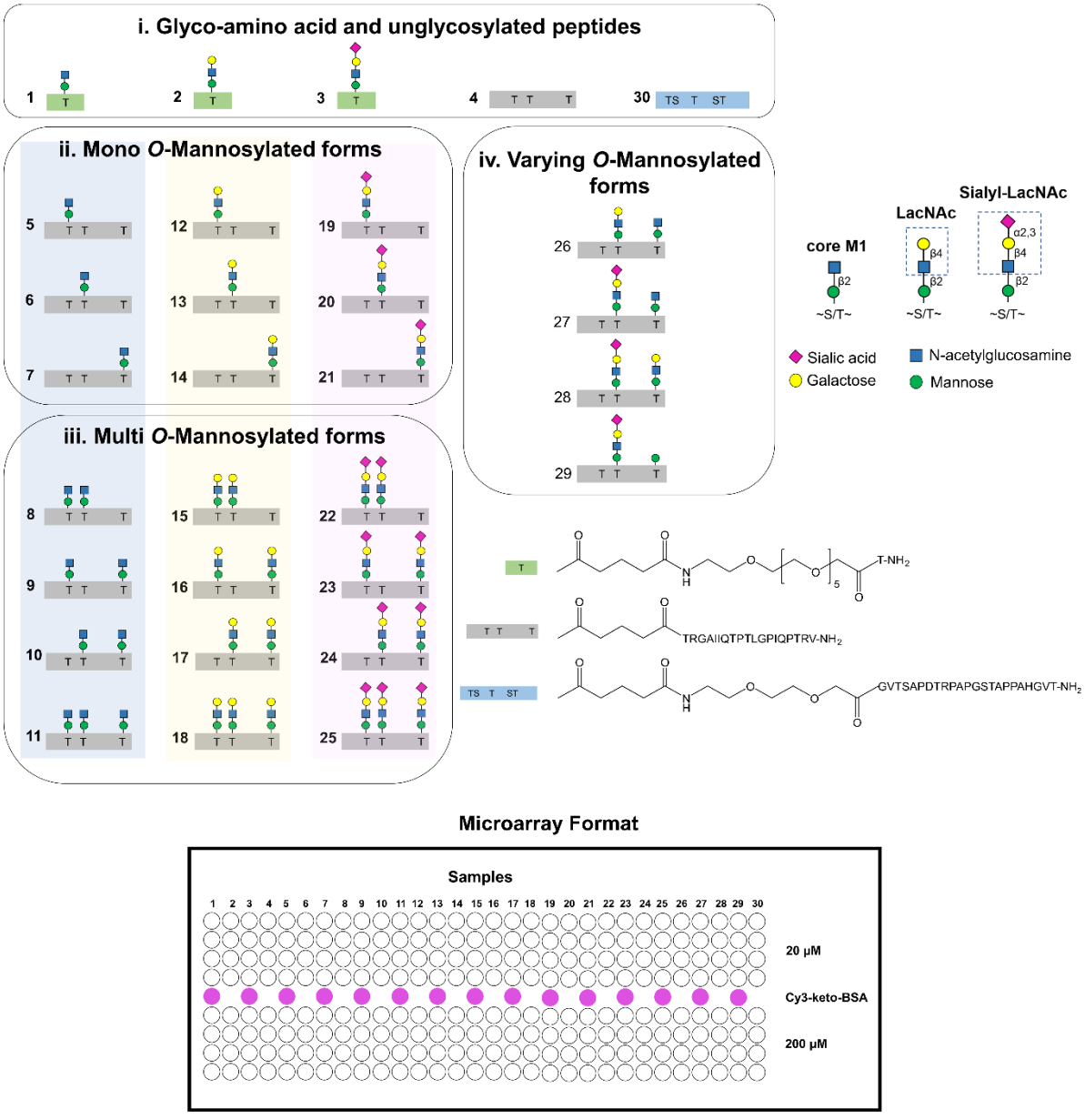

Figure 3. Versatile core M1 *O*-Mannosylated α-DG glycopeptide library utilizzed in microarray experiment.

domain of α-DG (^{372}TRGAIIQ**T**P**T**LGPIQP**T**RV^{390}) displaying *O*-Man core M1 and extended glycan structures (26, 34). This chemoenzymatically synthesized glycopeptide library can investigate various evaluation at the same time (Fig. 3): i) the individual contribution of glyco-amino acid and peptide scaffold to the lectin binding (**1-4**) and MUC1 peptide scaffold as negative control (**30**); ii) the impact of individual glycopeptide and position effect of mono core M1 (GlcNAcβ1-2Man, **5-7**) LacNAc-terminated core M1 (Galβ1-4GlcNAcβ1-2Man, **12-14**) and sialyl-LacNAc-terminated core M1 (NeuAcα2-3Galβ1-4GlcNAcβ1-2Man, **19-21**); and iii) the multivalency effect of bis- and tris-core M1 and extended structures (**8-11**, **15-18**, **22-25**); iv) the effect of different neighboring glycan structure (**26-29**). These 5-oxohexanoic acid N-terminus functionalized compounds were printed on a hydroxylamine-coated microarray chip in quadruplets. Then immobilized via glycoblotting method (35). To assure validity of this protocol, positive controls plant lectins *Concanavalia ensiformis agglutinin* (ConA) and *Glycine max* (soybean) agglutinin (SBA) were utilized as previously reported (26). Now the stage is set to determine the binding profile of the prepared glycoconjugates with galectin variants.

5.2.3. Binding profile of Gal-1 variants

The Gal-1 variants can bind to glycans of this test panel, so strong signals are recorded (Fig. 4 and Supplemental Figs. S1-S2). Similar to the wild type, Gal-1 variants maintained a preferential contact formation with core M1 LacNAc-terminated epitopes of α-DG irrespective of the modular design. The GG-linked homodimer (**(Gal-1)$_2$[GG]**) and tetramer (**(Gal-1)$_4$[GG]**) exhibited the same binding profile as the wild-type protein. They strongly interact with glyco-amino acid (**2-3**), unglycosylated peptide (**4**), and glycopeptides presenting mono core M1 (**5**). Also, glycopeptides terminated with LacNAc (**12-18, 25**) or sialyl-LacNAc (**19-25, 27-29**) were solid binders but not MUC1 peptide **30** (Fig. 4 and Supplemental Figs. S1-S2). No binding with compounds **6-11** was observed, specifically on 32 μg/mL lectin concentration,

differentiating the activity of parent protein under ambient oxidizing conditions (16). The rigid short linker like glycine-glycine fixates the homodimer Gal-1 prohibiting it from dissociation and prevents internal cysteine residue oxidation (36).

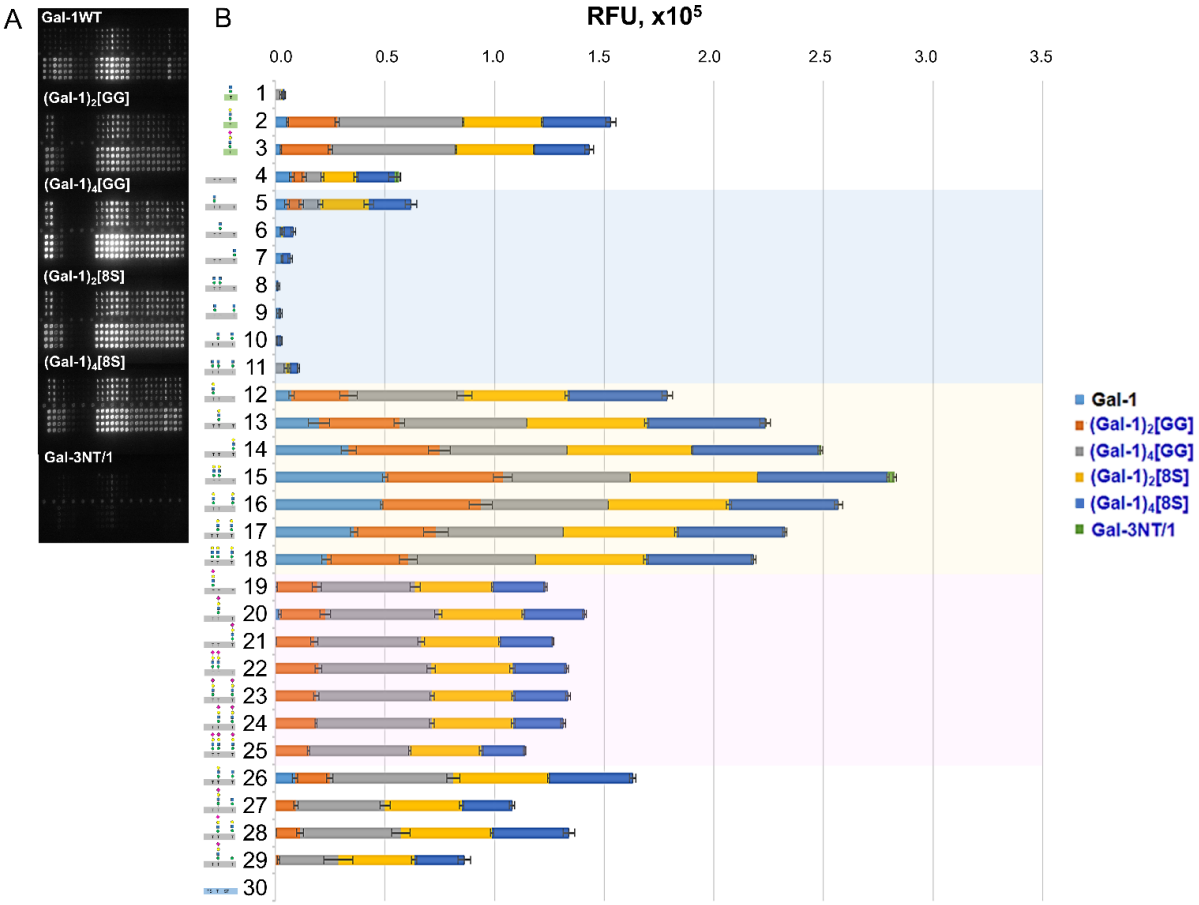

Figure 4. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with 10.0 μg/mL Gal-1 variants (**a**). Signal intensities stacked chart of 200 μM core M1 α-DG glycoconjugates with 3.20 μg/mL Gal-1 variants (**b**). The relative interaction of 20 and 200 μM core M1 α-DG glycoconjugates with 0.10 to 32.0 μg/mL of Gal-1 variants are in Supplementary Fig. S1 and S2, respectively.

Extending the linker length of Gal-1 variants (homodimer (**(Gal-1)$_2$[GG]**) and tetramer (**(Gal-1)$_4$[GG]**)) to 33 amino acids of Gal-8 (8S), yielding **(Gal-1)$_2$[8S]**) and **(Gal-1)$_4$[8S]**) respectively, demonstrated analogous binding capacity (Fig. 4 and Supplemental Figs. S1-S2). The presence of the linker resulted in enhanced binding of Gal-1 to the test compounds, and the length of the linker seemed to matter; the longer the size, the higher the binding activity (**(Gal-1)$_2$[GG]** vs. **(Gal-1)$_2$[8S]** and **(Gal-1)$_4$[GG] vs. Gal-1)$_4$[8S]**). The increase in linker

length allowed flexibility of the orientation and lateral movement of CRDs (37). In, addition, the **[8S]** linker perturbs the F-face of CRD possibly making it more accesible for binding thereby enabling more effective crosslinking of glycan ligands, especifically to core M1-terminated glycopeptides **6-11** (**(Gal-1)$_4$[GG]** vs. **(Gal-1)$_4$[8S]**, Supplemental Figs. S1-S2) (38). Furthermore, the increased number of CRDs per molecule did not change the binding properties relative to the parent homodimer, but enhanced the signal intensity (**(Gal-1)$_2$[GG]** vs. **(Gal-1)$_4$[GG]** and **(Gal-1)$_2$[8S]** vs. **(Gal-1)$_4$[8S]**). Furthermore, the increased number of CRDs per molecule did not change the binding properties but enhanced signal intensity relative to the parent homodimer were observed (**Gal-1WT** vs. **(Gal-1)$_4$[GG]** vs. **(Gal-1)$_4$[8S]**).

Bringing the Gal-1 CRD into the Gal-3-like design (**Gal-3NT/1**), markedly reduced signal intensities. Only at high concentrations (Fig. 4 and Supplemental Figs. S1-S2) this variant retained interaction with peptide (**4**), and glycopeptides terminated with core M1 (**5-7**) or LacNAc (**12-18,26,** and **18**). In this modular design, the Gal-1 CRD can use the N-terminal part of Gal-3 WT to oligomerize in concentration-dependent manner forming disorganized crosslinked complexes, similar to Gal-3WT (39). However, weak intensity signals were recorded, suggesting that Gal-1 CRD would prefer its wild type or be in tandem-repeat type variant form to bind to the surface-immobilized core M1 α-DG ligands.

5.2.4. Binding profile of Gal-3 variants

The wild-type Gal-3 interacted weakly with LacNAc-terminated glycopeptides **12-18** (Fig. 5 and Supplemental Figs S3-S4). The removal of the collagenous N-terminal tail (**trGal-3**) resulted in a weak enhancement of binding for LacNAc-terminated (**12-18**, **26**) and sialyl-LacNAc (**3**, **19-25**, **27**) core M1 α-DG glycoconjugates (Fig. 5 and Supplemental Figs S3-S4). The loss of N-terminal tail in Gal-3 disables it from forming dimers or higher-order oligomers (40). In addition, unlike the prototype Gal-1 and Gal-2 which can form dimers via hydrophobic

interactions between N- and C-terminal residues of two monomeric units, the **trGal-3** cannot form dimers by the same interactions (41). Thus, the slight increase in affinity suggests that the CRD only modular architecture is desirable for Gal-3 so that it can interact with the mobilized ligands. The results here further reflected the preference of Gal-3 CRD for reducing-end polyLacNAc structures, not on the mono-terminal LacNAc structures as displayed by the α-DG glycoconjugates (42, 43).

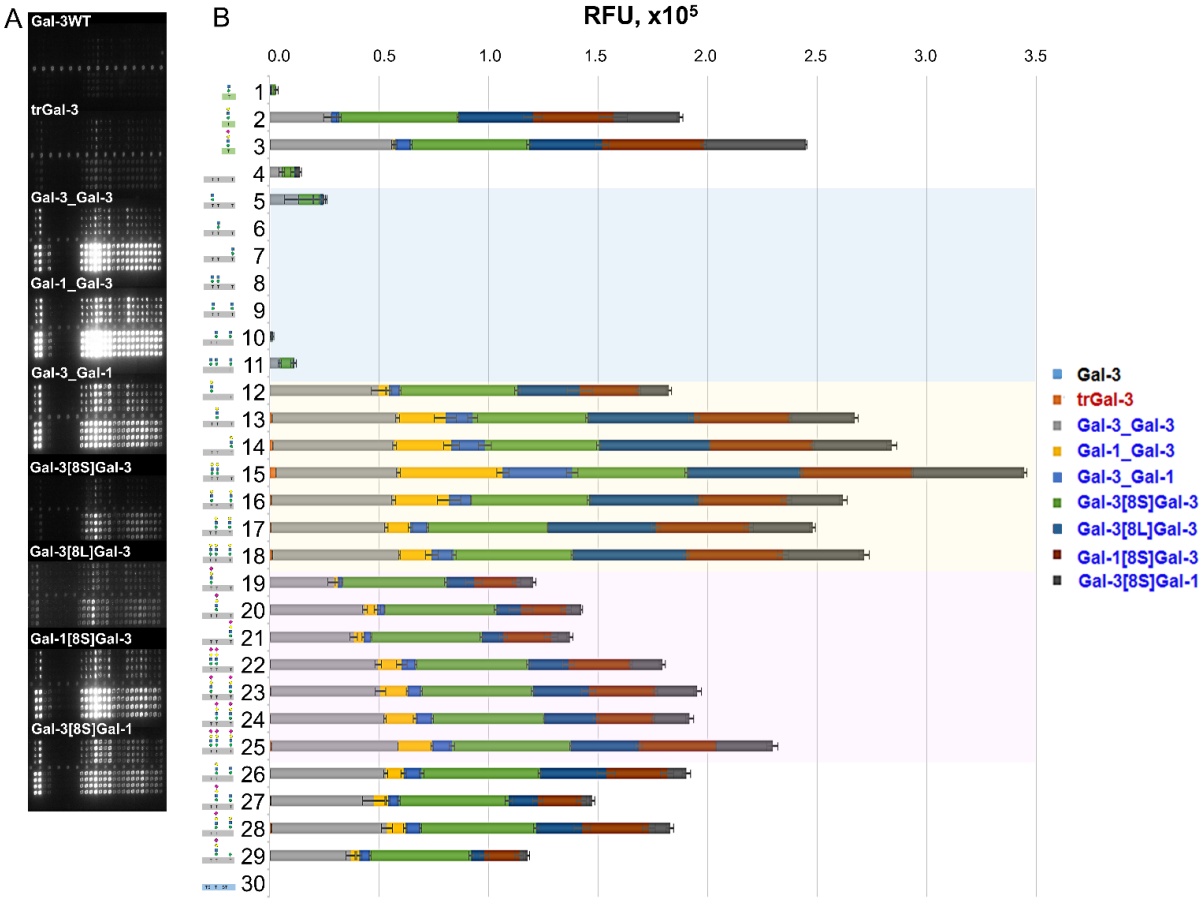

Figure 5. (A.) Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with **10.0 µg/mL** Gal-3 variants. (B.) Signal intensities stacked chart of **200 µM** core M1 α-DG glycoconjugates with **3.20 µg/mL** Gal-3 variants (b). The relative interaction of 20 and 200 µM core M1 α-DG glycoconjugates with 0.10 to 32.0 µg/mL of Gal-3 variants are in Supplementary Fig. S3 and S4, respectively.

The homodimer (**Gal-3_Gal-3**) and heterodimers (**Gal-1_Gal-3** and **Gal-3_Gal-1**) prototype variants are highly associated with LacNAc-terminated (**2** and **12-19**), and its α2,3-sialylated derivatives (**3** and **19-25** and **27-29**) (Fig. 5 and Supplemental Figs S3-S4). Very weak interaction was recorded for core M1 glyco-amino acid **1** and unglycosylated peptide **4**.

Conjugation of **1** to the glycosylation site of **4** (**5-11**) did not result in enhanced binding of these variants. The spatial orientation of Gal-1/-3 in heterodimer prototype variants (**Gal-1_Gal-3** vs. **Gal-3_Gal-1**) affected the binding interaction to α2,3-sialylated core M1 α-DG glycopeptides (**19-25** and **27-29**).

Transforming the homodimeric Gal-3 (**Gal-3_Gal-3**) to tandem-repeat types **Gal-3[8S]Gal-3** and **Gal-3[8L]Gal-3 negatively influenced its interaction with the test glycoconjugates** (Fig. 5 and Supplemental Figs. S3-S4). The tandem-repeat types variants both exhibited a decrease in affinity on LacNAc- and sialyl-LacNAc terminated core M1 glycoconjugates (**Gal-3_Gal-3** vs. **Gal-3[8S]Gal-3** vs. **Gal-3[8L]Gal-3**). The 8L-linked Gal-3 (**Gal-3[8L]Gal-3**), but not 8S-linked Gal-3 (**Gal-3[8S]Gal-3**), interacted with glyco-amino acids **1-3**, unglycosylated α-DG peptide **4**, and core M1 glycopeptide **5**. In the case of 8S-linked Gal-1/Gal-3 heterodimers, the presence of an 8S-linker also negatively influenced the binding activity of this variant against sialyl-LacNAc- presenting α-DG glycopeptides **19-25** and **27-29** (**Gal-1_Gal-3** vs. **Gal-1[8S]Gal-3** and **Gal-3_Gal-1** vs. **Gal-3[8S]Gal-1**) (Fig. 5 and Supplemental Figs. S3-S4). The consequence of Gal-1/Gal-3 CRD spatial orientation is more prominent on the tandem-repeat type than on prototype variants, specifically in low ligand concentration.

Overall, the results suggest that removing the N-terminal tail of Gal-3 resulted in increased affinity to α-DG core M1 glycoconjugates. Converting the chimera-type parent protein into directly conjugated prototype-like homo- (Gal-3) and heterodimer (Gal-1/Gal-3) variants led to a positive binding. Incorporating linkers into Gal-3 prototype variants generally led to decreased affinity. Furthermore, the order of modular assembly of Gal-1/ Gal-3 CRD from N- to C-terminal had an effect. These observed activities reflected the binding properties, obtained K_d values, of these Gal-3 variants with neuroblastoma cells (28).

5.2.5. Binding profile of tandem-repeat type (Gal-4 and -9) variants

The wild-type Gal-4 interacts with peptide **4**, mono-core M1 substituted (**5-7**), and LacNAc-terminated core M1 glycopeptides (**12-18**) (Fig. 6 and Supplemental Figs. S5-S6). Reducing the size of the peptide linker (**Gal-4V**) and conversion to a prototype (**Gal-4P**) reduced the level of reactivity to the identical ligands (**4**, **12-18**) as compared to **Gal-4WT** (Fig. 6 and Supplemental Figs. S5-S6). The transformation of the parent protein to a Gal-3-like protein (**Gal-3NT/4C**) abolished its binding activity to the core M1 glycopeptides (Fig. 6 and Supplemental Figs. S5-S6). These results demonstrated that reduced linker length or its complete removal, i.e. changing Gal-4WT to a prototype-like structure, decreased the lectin's

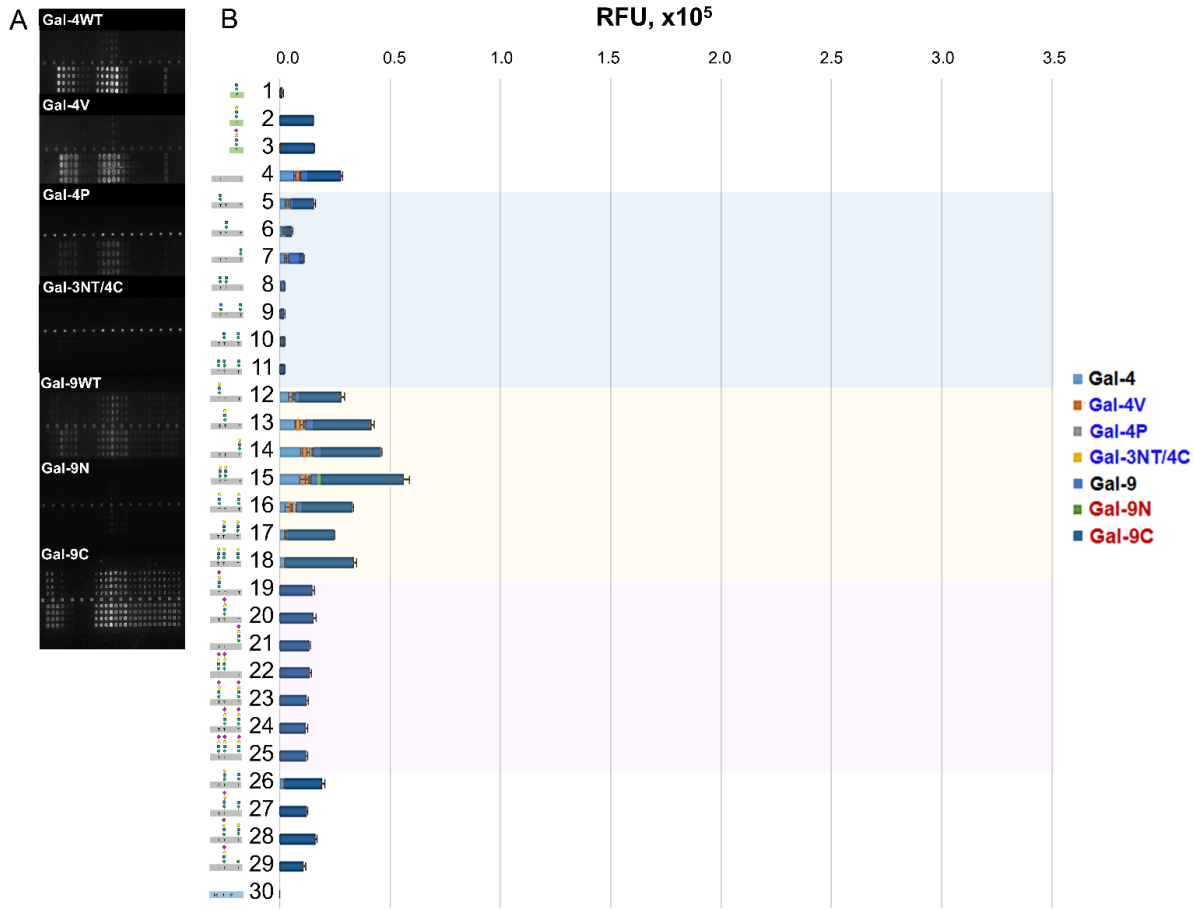

Figure 6. (A.) Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with **10.0 μg/mL** Gal-4 and -9 variants. (B.) Signal intensities stacked chart of **200 μM** core M1 α-DG glycoconjugates with **3.20 μg/mL** Gal-4 and -9 variants (b). The relative interaction of 20 and 200 μM core M1 α-DG glycoconjugates with 0.10 to 32.0 μg/mL of Gal-4 and -9 variants are in Supplementary Fig. S5–S6 and S7–S8, respectively.

susceptibility to the test compounds.

Moreover, presenting tandem-repeat-type Gal-4 CRDs in a chimera-type abolishes its binding activity towards the surfaced immobilized ligands (**Gal-4WT>Gal-4V>Gal-4P>Gal3NT/4C**). The linker is not just the connector of the CRDs but also influences the lectin affinity because it provides spatial arrangement flexibility of the CRDs, as observed with Gal-4 interaction with glycoclusters (31, 44).

The wild-type Gal-9 revealed moderate affinity to unglycosylated peptide **4,** core M1-presenting glycopeptides (**5- 11**) and LacNAc-terminated compounds (**12-18**) (Fig. 6 and Supplemental Figs. S7-S8). These observations were modulated by the sugar epitope positioning and density. The α2,3 sialylation of the LacNAc units (**19-25, 27-29**) impaired the binding of Gal-9 to the glycoconjugates. The N-terminal CRD (**Gal-9N**) is bound to a few glycoconjugates, here only compounds **15** and **18** (Supplemental Figs. S7-S8). On the other hand, the **Gal-9C** revealed an analogous binding profile with Gal-9WT, having increased interaction at lower lectin concentration and enhanced affinity to sialyl-LacNAc-terminated glycopeptides (**19-25, 27-29**) (Supplemental Figs. S7-S8). These results suggest the difference in binding activity of the C- and N-terminal CRD of Gal-9, where **Gal-9C** likely contributes to the binding of its wild-type form and that the presence of linker and N-terminal CRD affected its overall affinity.

5.2.6. Gal-1 variants trans-bridge α-DG core M1 glycopeptides and laminins in microarray

Gal-1 is a promising candidate in reducing muscular dystrophy disease pathology; however, the exact mechanisms are still unknown (25, 45, 46). Previously, we demonstrated the crosslinking capabilities of wild-type Gal-1 between various laminins and α-DG peptide

and core M1 glycopeptides in microarray (26). We attribute this to the protein-peptide interaction of Gal-1 F-face with α-DG peptide **4** and protein-carbohydrate interaction of Gal-1 S-face with polyLacNAc-terminated *N*-glycans presented by laminin. This finding proposes a possible mechanism by which Gal-1 can ameliorate several forms of congenital muscular dystrophy where α-DG is highly hypoglycosylated. As decribed above, Gal-1 variants (except **Gal-3NT/1**) and Gal-1/Gal-3 heterodimer-tandem-repeat type variants (**Gal-1[8S]Gal-3**, **Gal-3[8S]Gal-1**) revealed enhanced affinity to the prepared core M1 glycoconjugates of α-DG, Fig. 4 and 5. For possible translational therapeutic applications in treating α-dystroglycanopathy by *trans*-bridging laminin and core M1 glycans of α-DG, we selected Gal-1 presenting variants (**(Gal- 1)$_2$[8S], Gal-1[8S]Gal-3**, **Gal-3[8S]Gal-1**, **(Gal-1)$_4$[8S]**, and **Gal-3[8S]Gal-3** (negative control)).

The parent Gal-1 exhibited high *trans*-bridging activity between laminin-111, -121, -211, and -221 (but little -511) and α-DG peptide (**4**) and core M1 glycopeptides (**5-11**) in glycan position- and density-dependent manner, Fig. 7. Presenting Gal-1 into a tandem-repeat type having an 8S-linker (**(Gal-1)$_2$[8S]**) maintained its crosslinking activity with laminin and α-DG peptide **4**, but attenuated *trans*-bridging with core M1 glycopeptides (**5-11**) were observed. Interestingly, a noticeable increase in *trans*-bridging towards LacNAc-terminated (**12–18, 26**) and sialyl-LacNAc-terminated (**19–25**, **27–29**) ~~were~~ was demonstrated by **(Gal-1)$_2$[8S]** that is attributed to the high affinity of this Gal-1 variant to the said ligands compared to **Gal-1WT**, Fig 7. This difference in crosslinking activity is associated with the Gal-8 linker, 33 aa (8S), between two Gal-1 monomers providing Gal-8-like properties to Gal-1. Gal-8 interacts with a receptor peptide fragment of NDP52 activating antibacterial autophagy (47). The X-ray structure of this interaction revealed that the C- and N-terminal CRDs could be associated with the ligand in back-to-side orientation, forming dimeric structures, thus possessing four carbohydrate recognition domains available for binding (44, 48). Due to this, **(Gal-1)$_2$[8S]** is

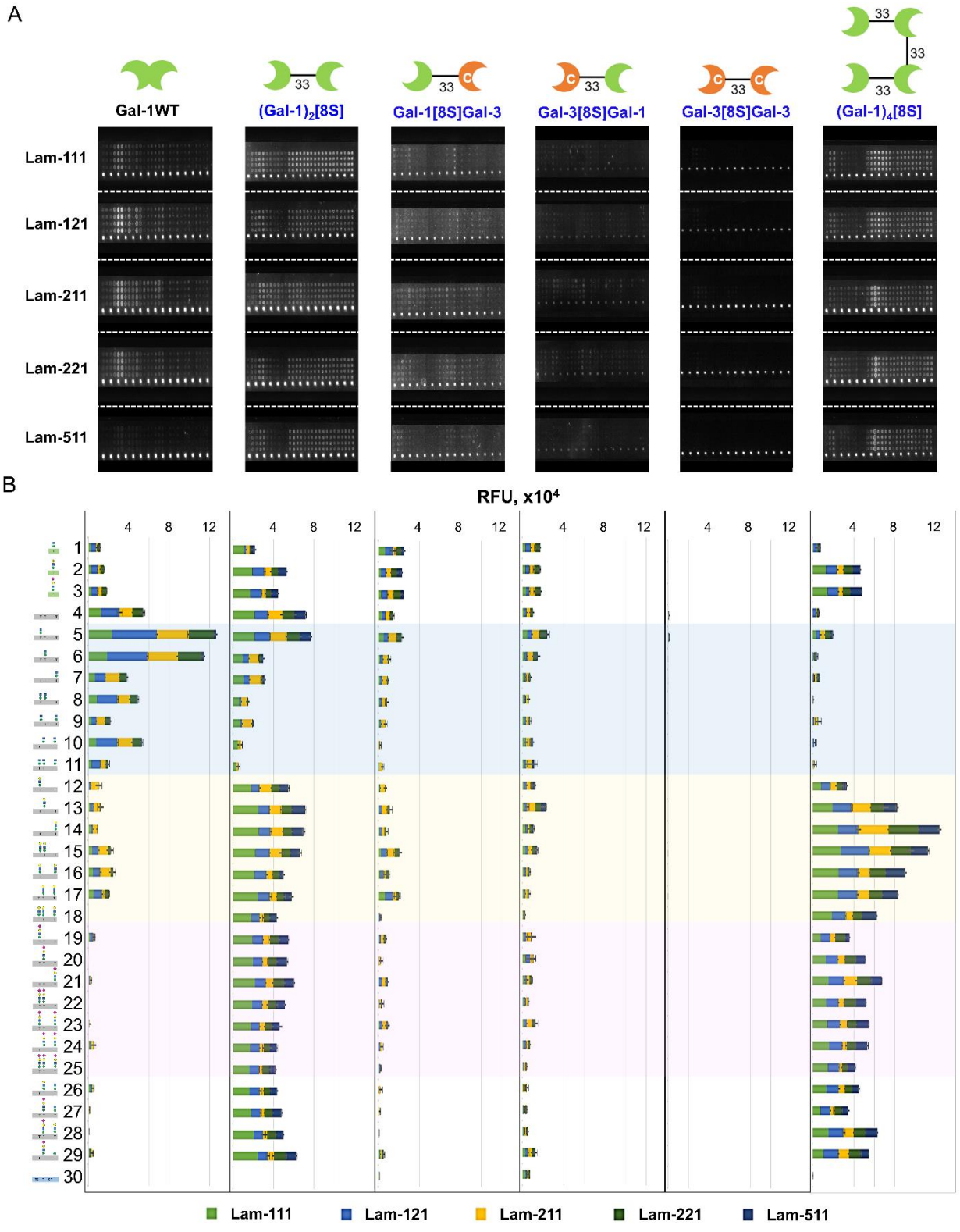

Figure 7. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip (**a**) and stacked chart of signal intensities (**b**) taken after treatment with various 32.0 μg/mL **laminin–Gal-1 variant** solution for 30 min. The relative interaction of 200 μM core M1 α-DG glycoconjugates with 0.10 to 32.0 μg/mL of various **laminin–Gal-1 variant** solution are in Supplementary Fig. S9–S14.

expected to demonstrate improved crosslinking activity towards the LacNAc- and sialyl-LacNAc-terminated core M1 glycopeptides of α-DG and laminin than **Gal-1WT**. Moreover, **(Gal-1)$_2$[8S]** can *trans*-bridge core M1 glycoconjugates of α-DG with Lam-511, which ~~is~~ was not exhibited by the wild-type protein. Overall, these results suggest that the *trans*-bridging activities of **(Gal-1)$_2$[8S]** occurs via F-Face (Gal-1-α-DGpeptide)-to-S-face (Gal-1-LacNAc-terminated glycans of laminin) and S-face (Gal-1-α-DG LacNAc-terminated glycopeptides)-to-S-face (Gal-1-LacNAc-terminated glycans of laminin) interactions, Fig. 8.

On the other hand, converting **(Gal-1)$_2$[8S]** to a hetero-tandem-repeated type 8S-linked Gal-1/-3 variants (**Gal-1[8S]Gal-3** and **Gal-3[8S]Gal-1**) led to a considerable reduction in the *trans*-bridging activity on the surface-immobilized ligands and in solution laminins, Fig. 7. The **trGal-3** showed very weak interaction with the prepared ligands compared to **Gal-1WT**. Incorporating the Gal-3 CRD with Gal-1 via 8S-linker (**Gal-1[8S]Gal-3** or **Gal-3[8S]Gal-1**) resulted in increased binding activity is credited to the high- affinity Gal-1 CRD with the test α-DG core M1 glycoconjugates, Fig. 5 and Supplemental Figs S3-S4. Gal-3 is also known to bind to laminin through its numerous poly-LacNAc structures (49). It can bridge neutrophils to laminin *in vitro,* which is dependent on both CRD and the N-terminal terminal tail region (50). However, **Gal-1[8S]Gal-3** and **Gal-3[8S]Gal-1** failed to crosslink effectively core M1 glycoconjugates of α-DG to laminin. These results suggest that, at least two Gal-1 CRD is needed to effectively trans-bridge α-DG glycopeptides with laminins. As expected, the **Gal-3[8S]Gal-3** variant did not show any trans-bridging activity to any of the prepared α-DG ligands and laminins, Fig. 7.

The **(Gal-1)$_4$[8S]** exhibited high crosslinking ability towards LacNAc-presenting α-DG core M1 glycoconjugates than **Gal-1WT** and **(Gal-1)$_2$[8S]**, while a substantial decrease in signal intensity is observed in the α-DG peptide **4** and core M1-terminated glycopeptides (**5–11**), Fig. 7. Given the increased number of Gal-1 CRD sites per protein, the enhancement in

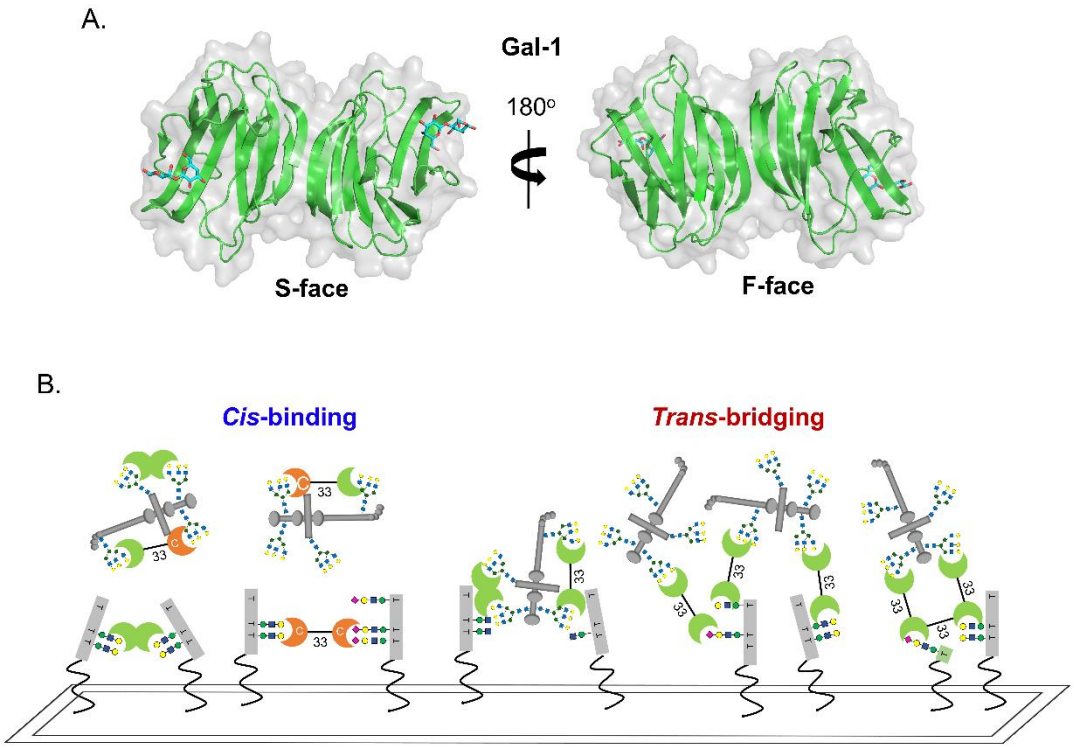

Figure 8. (A.) The X-ray crystal structure of Gal-1 (PDB access code: 1GZW) showing the S- and F-face of the CRD . (B.) The *cis*- and *trans*-bridging activity activities of Gal-1 variants with core M1 α-DG and laminin *in situ*.

trans-bridging of **(Gal-1)$_4$[8S]** is anticipated. In this case, the high activity towards LacNAc-terminated core M1 glycoconjugate indicates that S-Face (Gal-1-α-DG LacNAc/sialyl-LacNAc-terminated glycopeptides) to S-Face(Gal-1 - LacNAc-terminated glyco-peptides of laminin) *trans*-bridging interaction with laminins is involved Fig. 8.

5.3. Conclusion

Changes in galectin architecture such as i) turning a non-covalently associated homodimer into a covalently linked homodi- or tetramer or a heterodimer, ii) bringing a CRD into a different structural context or iii) reducing the linker length of a tandem-repeat-type galectin can modulate the binding towards core M1 α-DG glycoconjugates. These classes of glycans disclosed that Gal-1's CRD maintains its binding pattern (with LacNAc and its

oligomers) irrespective of the tested alterations of the protein architecture when interacting with surface-immobilized glycans in an array. In the case of Gal-3, removing the N-terminal tail and conversion to mono/hetero tandem-repeat type can enhance its affinity to the prepared ligands, and the presence of peptide linker between the CRDs influenced the activity. For Gal-4, maintaining the full-length linker is crucial, and its transformation to prototype and chimera type is undesirable for binding. On the other hand, the C-terminal Gal-9 CRD mainly contributed to the overall binding of the wild-type protein to core M1 α-DG core glycoconjugates.

Furthermore, the Gal-1 variants **(Gal-1)$_2$[8S]** and **(Gal-1)$_4$[8S]** can effectively *trans*-bridge core M1 α-DG glycoconjugates and laminins (111, -121, -211, -221, and -511) compared to Gal-1 wild-type. This enhancement of crosslinking activity is attributed to the additional peptide linker between two Gal-1 CRD and the increased number of Gal-1 CRD sites per protein. Similar to Gal-1WT, these Gal-1 variants are also good potential candidates for treating α-dystroglycanopathy. However, further tests must be carried out to validate this proposal.

Overall, this experimental setup demonstrated that the alteration of the galectin structures can give additional insights into the preferential modular architecture and binding behavior of this lectin family towards specific ligands. In addition, rational protein engineering is a useful tool in redesigning lectins with possibly higher therapeutic potentials than their wild-type counterpart. Here, Gal-1 has proof-of-principle character.

5.4. Experimental section

5.4.1. Materials

Commercially available solvents and reagents were purchased from Sigma-Aldrich (USA), Tokyo Chemical Industry (Tokyo, Japan), Wako Pure Chemical Industries (Osaka, Japan), Kokusan Chemical (Tokyo, Japan), or Watanabe chemical (Osaka, Japan) and used without purification, unless stated. Laminins (111, 121, 211, 221, and 511) and Cy3 NHS antibody labeling kit were procured from Biolamina and BroadPharm, respectively.

5.4.2. Methods

Production of labeled galectin variants (4, 28, 51)

Galectins were prepared by recombinant expression using the *E.coli* strain BL21 (DE) pLysS and the pGEMEX-1 vector (Promega, Walldorf, Germany). Galectin expressing bacteria were grown at 37°C until an OD_{600} of 0.6-0.8. Expression of i) Gal-1, $(Gal\text{-}1)_2$[GG], $(Gal\text{-}1)_4$[GG], $(Gal\text{-}1)_2$[8S], and $(Gal\text{-}1)_4$[8S] was induced with 100 µM isopropyl ß-D-1-thiogalactopyranoside (IPTG) at 37°C, ii) Gal-3 (400 mM IPTG), trGal-3, Gal-3—Gal-3, Gal-3—Gal-1, Gal-1—Gal-3, Gal-3—8S—Gal-3, Gal-3—8L—Gal-3, Gal-3—8S—Gal-1, Gal-1—8S—Gal-3, Gal-9N, and Gal-9C with 100 µM IPTG at 22°C, and iii) Gal-4, Gal-4V, Gal-4P, and Gal-3NT/4C at 30°C with 75 µM IPTG. Induced bacteria were grown for 16 h. Proteins were purified after cell lysis by sonification (three times, each stroke 1 min) through affinity chromatography on a home-made lactose-sepharose resin. Afterward, bound proteins were eluted from the resin using 20 mM PBS, pH 7.2, containing 50 mM lactose and 20 mM iodoacetamide. PBS was exchanged employing a PD10 column to 10 mM sodium carbonate buffer, pH 8.5. The proteins (2 -3 mg/mL) were directly conjugated in the dark and in the presence of activity-preserving 20 mM lactose to NHS-ester Alexa 555 fluorescent dye at 25°C for 4 h. Unbound dye was removed by gel filtration with a Sephadex G-25 column. Protein

purity was checked by gel electrophoresis and Western blotting. Activity were verified by solid-phase assays, flow cytometry, and hemagglutination. Labeled proteins were lyophilized in aliquots and stored at -20°C until reconstitution. A 1 mg/mL stock probed galectins were prepared in 1x PBS (pH 7.40), containing 1% (w/v) BSA, 0.09% (w/v) NaN_3, and 50% (w/v) glycerol and stored at -20^{o}C till use.

Development of α-dystroglycan core M1 (glyco)peptide library (14, 34)

The *O*-Man core m1 α-DG glyco-amino acids (**1**), peptide (**4**), glycopeptides (**5-11**), and MUC1 peptide (**30**) were synthesized manually by microwave-assisted solid-phase synthesis by using Fmoc-amino acids (Novabiochem), Fmoc(Ac_3GlcNAcβ1→2Ac_3Manα)Thr (Medicinal Chemistry Pharmaceuticals, Sapporo, Japan), and H-Rink Amide ChemMatrix® (0.48 mmol/g, 24 μmol) resin. The resin was swollen with CH_2Cl_2 in a polypropylene tube equipped with a filter (LibraTube, Hipep Laboratories, Kyoto, Japan) for 1 h at room temperature. The protected Fmoc-amino acid (4.0 equiv) was pre-activated by treating with HBTU (4.0 equiv), HOBt (4.0 equiv), and DIEA (6.0 equiv) in DMF (455 μL) for 9 min under microwave irradiation (Green Motif 1 microwave synthesis reactor, IDX Corp, Japan) and then attached to the resin. In every step, the N-fluorene-9-ylmethoxycarbonyl (Fmoc) groups at N-terminal were removed by 20% piperidine in DMF (1 mL) for 3 min under microwave irradiation. All coupling reactions were done for 10 min, and removed solvents using PP syringes fitted with a porous disk. For glycosylated amino acid, Fmoc-Thr(Ac_3GlcNAcβ1→2Manα1)-OH (1.2 equiv) was treated with PyBOP (1.2 equiv), HOAt (1.2 equiv) and DIEA (3.0 equiv) in DMF (275 μL) subjected to MW irradiation for 9 min at 50°C. After which, PyBOP-HOAt (1.2 equiv) was added and allowed to react for another 9 min. As the final synthesis step, 5-oxohexanoic acid (3 equiv) was introduced at the N-terminus of each glycopeptide resin, according to the above coupling procedure for Fmoc-amino acids. Removal of side-chain protecting groups and cleavage of glycopeptide from the resin occurred in parallel

by treatment with 95% aqueous TFA (1 mL) for 1 h at ambient temperature. The crude peptides and glycopeptides were precipitated in a cold water bath by tert-butyl methyl ether (5 mL). After that, the solution was centrifuged at 3000 rpm for 1 min, and the supernatant was carefully removed. The precipitate was dissolved in milli-Q water (5 mL), and lyophilized. Then, deacetylation of the glycan moiety was done by dissolving the lyophilized material in methanol The pH was adjusted at 12.5 with dropwise addition of 1M NaOH, and the solution was stirred at room temperature for 1 h. After the deprotection, the solution was neutralized with 20% AcOH in methanol, and a flow of nitrogen gas displaced the solvent. The crude peptides and glycopeptides were purified by RP-HPLC, using a preparative C18-reversed-phase column (Intersil ODS-3 10×250 mm) on HPLC (HITACHI, Japan) equipped with L7150 pump, at a flow rate of 5 mL per min monitored by UV detector at 220 nm at room temperature. Eluent A was distilled water containing 0.1% TFA, and eluent B was acetonitrile containing 0.1% TFA. Each product was analyzed by Ultraflex MALDI-TOFMS (Bruker Daltonics, Germany) using DHB as a matrix.

Galactosylation of compounds 1, 5–11 in a 24 h incubation step yielded compounds 2, 12–18. The reaction's 50 mM HEPES buffer (pH 7.0) contains 10 mM $MnCl_2$, 0.1 wt% BSA, galactosyltransferase from bovine milk (Sigma Aldrich), and UDP-Gal (Yamasa Corporation, Chiba, Japan). Subsequently, compounds **2, 12**–**18** were sialylated using α2,3-sialyltransferase from *Pasteurella multocida* (Sigma Aldrich) and CMP-NANA (Yamasa Corporation, Chiba, Japan) in 50 mM Tris buffer (pH 6.5) and 500 mM NaCl incubated for 36 h to yield compounds **3, 19**–**25**. Compounds **26**–**29** were synthesized by solid-phase synthesis and enzymatic sugar elongations, as described earlier. Each product was purified by a RP-HPLC, as described above, with an appropriate solvent system and identified by MALDI-TOFMS.

Lectin Microarray Binding Experiment (14, 26, 52)

The AO/PC-copolymer microarray slides (Sumitomo Bakelite Co., Ltd., Tokyo, Japan) were deprotected using 2N HCl overnight at rt, rinsed with MilliQ H_2O, and dried by centrifugation. The test compounds were robotically printed in quadruplets at two concentrations (20 and 200 µM) in 25 mM AcOH-Pyr (pH 5.0), 0.0025% (w/v) Triton X-100 using an Arduino-based CNC machine handcrafted robot. A cyanin3-keto-BSA (Cy3-keto-BSA) at 25 µg/mL was also printed on the slide as a grid. Subsequently, the slide was incubated at 80°C for 1 h to complete the oxime bond formation. Washed once with Milli-Q H_2O and dried by centrifugation at 2,000 rpm for 2 min.

A silicon rubber sheet with six chambers was attached to the printed slide. Next, slide was pretreated with reaction buffer {Phosphate-Buffered Saline solution (PBS, 1X) [10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2.7 mM KCl, 137 mM NaCl] pH7.4 containing 0.05% (v/v) Tween-20} for 15 min and dried by centrifugation. The plant lectin or galectin solutions in PBS were prepared and maintained in a cold ice bath before use. A cover glass was then set on each chamber, and 12 µL of 0.10 µg/mL lectin solution in reaction buffer was added through the gap of the slide and cover. After 1 h of incubation with the lectin solution at rt in a humidified chamber, the slide was rinsed with washing buffer, and fluorescence intensity was measured with GlycoStation System (GlycoStation Reader 1200, GlycoTechnica Ltd., Yokohama, Japan). To determine additional interactions at higher Gal concentrations, the reaction buffer solution was carefully removed and replaced with the next lectin test concentration, incubated for 5 min at rt, washed with reaction buffer, and obtained fluorescence intensity. This step was repeated until all the chosen test concentrations were completed (0.32-, 1.00-, 3.20-, 10.0-, and 32.0 µg/mL).

Images of slides were captured in the presence of reaction buffer. The fluorescence intensities obtained were analyzed using ArrayVision software V8.0 (GE Healthcare, Tokyo, Japan). The background correction was applied to get the net intensity. The average relative fluorescence unit (RFU) was plotted as a bar graph and error bars being the standard deviation utilizing Microsoft Excel.

Laminin–Gal-1 Binding Assay (26)

The printed test compounds were added with 12 µL of premixed Cy3-laminin (Biolamina and BroadPharm) and unlabeled-Gal-1(variants) solutions with the final concentration of 0.10 µg/mL and incubated for 1 hr in a humidified chamber. The slide was rinsed with washing buffer and fluorescence intensity was measured with GlycoStation System, as described above. To determine additional interactions at higher laminin-Gal-1 concentrations, we carefully removed the reaction buffer and replaced it with the next test concentration, incubated for 5 min at rt, washed with reaction buffer, and obtained fluorescence intensity. This step was repeated until all the chosen test concentrations were completed (0.32-, 1.00-, 3.20-, 10.0-, and 32.0 µg/mL). For weak interaction, incubation with 32.0 µg mL^{-1} of laminin-Gal-1 solution was extended for 30 min. Slide images were captured and analyzed as described above.

5.4.3 Supplementary information

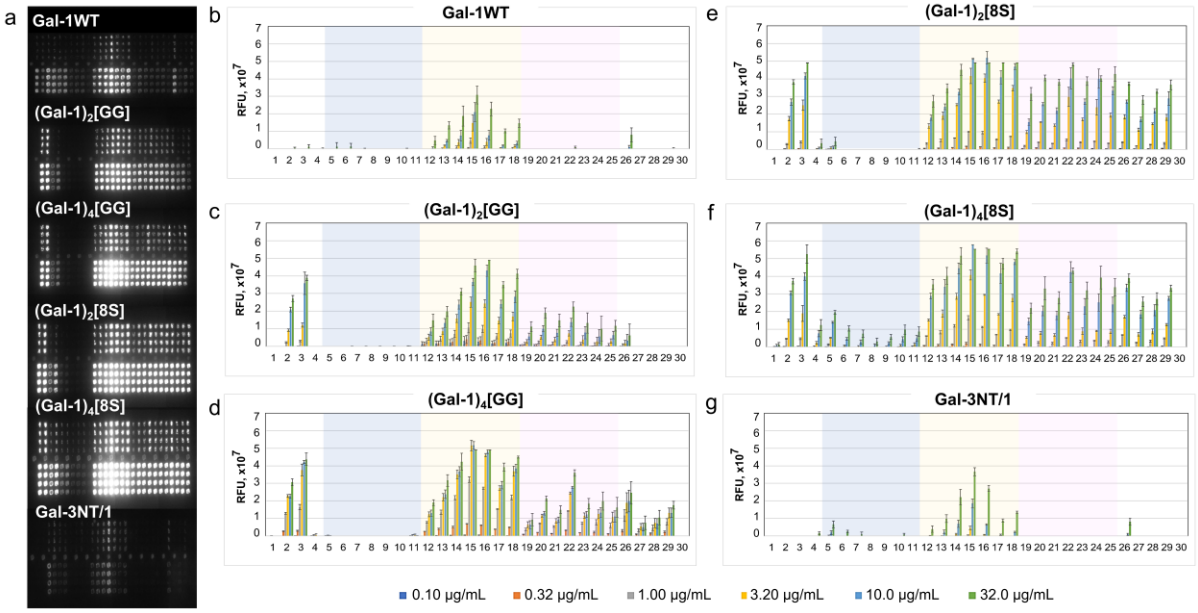

Figure S1. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with 10.0 µg/mL galectin solution (**a**) and relative interaction of **20 µM** core M1 α-DG glycoconjugates with 0.10 µg/mL to 32.0 µg/mL Gal-1WT (**b**), $(Gal\text{-}1)_2GG$ (**c**), $(Gal\text{-}1)_4GG$ (**d**), $(Gal\text{-}1)_28S$ (**e**), $(Gal\text{-}1)_48S$ (**f**), and Gal-3NT/1 (**g**).

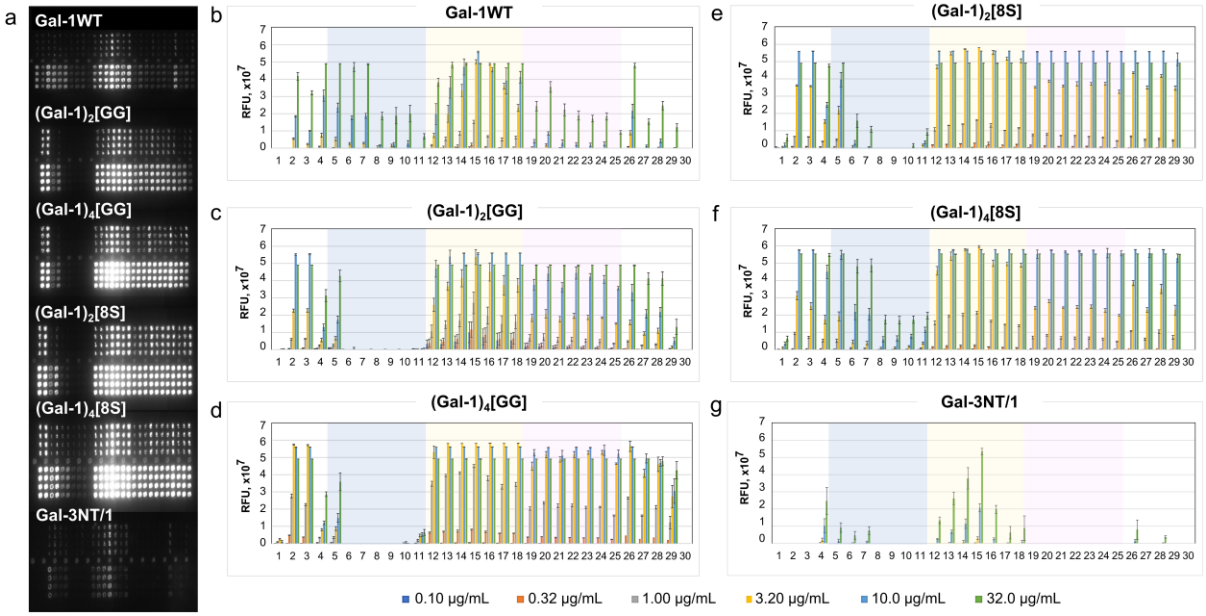

Figure S2. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with 10.0 µg/mL galectin solution (**a**) and relative interaction of **200 µM** core M1 α-DG glycoconjugates with 0.10 µg/mL to 32.0 µg/mL Gal-1WT (**b**), $(Gal\text{-}1)_2$GG (**c**), $(Gal\text{-}1)_4$GG (**d**), $(Gal\text{-}1)_2$8S (**e**), $(Gal\text{-}1)_4$8S (**f**), and Gal-3NT/1 (**g**).

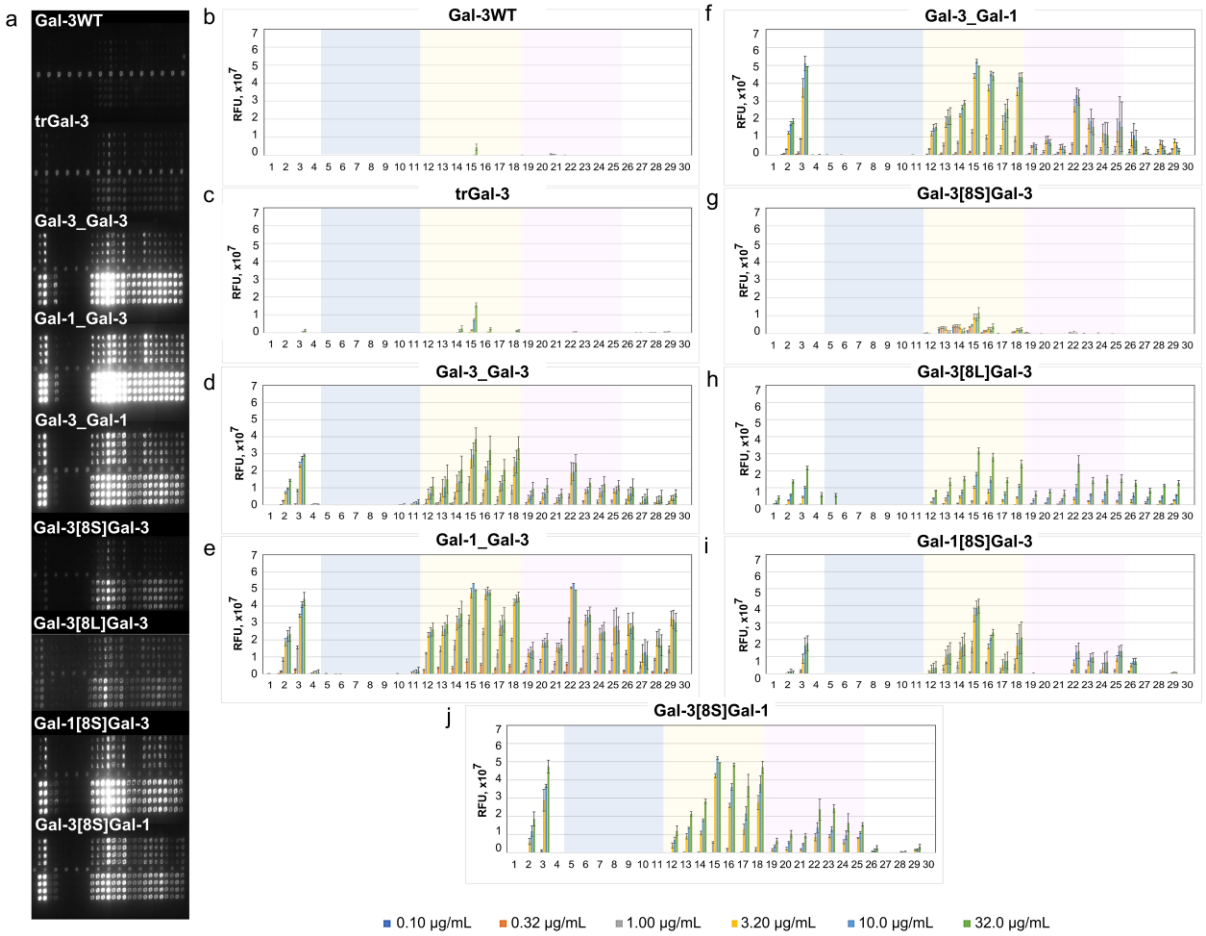

Figure S3. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with 10.0 µg/mL galectin solution (**a**) and relative interaction of **20 µM** core M1 α-DG glycoconjugates with 0.10 µg/mL to 32.0 µg/mL Gal-3WT (**b**), trGal-3 (**c**), Gal-3_Gal-3 (**d**), Gal-1_Gal-3 (**e**), Gal-3_Gal-1 (**f**), Gal-3[8S]Gal-3 (**g**), Gal-3[8L]Gal-3 (**h**), Gal-1[8S]Gal-3 (**i**), and Gal-3[8S]Gal-1 (**j**).

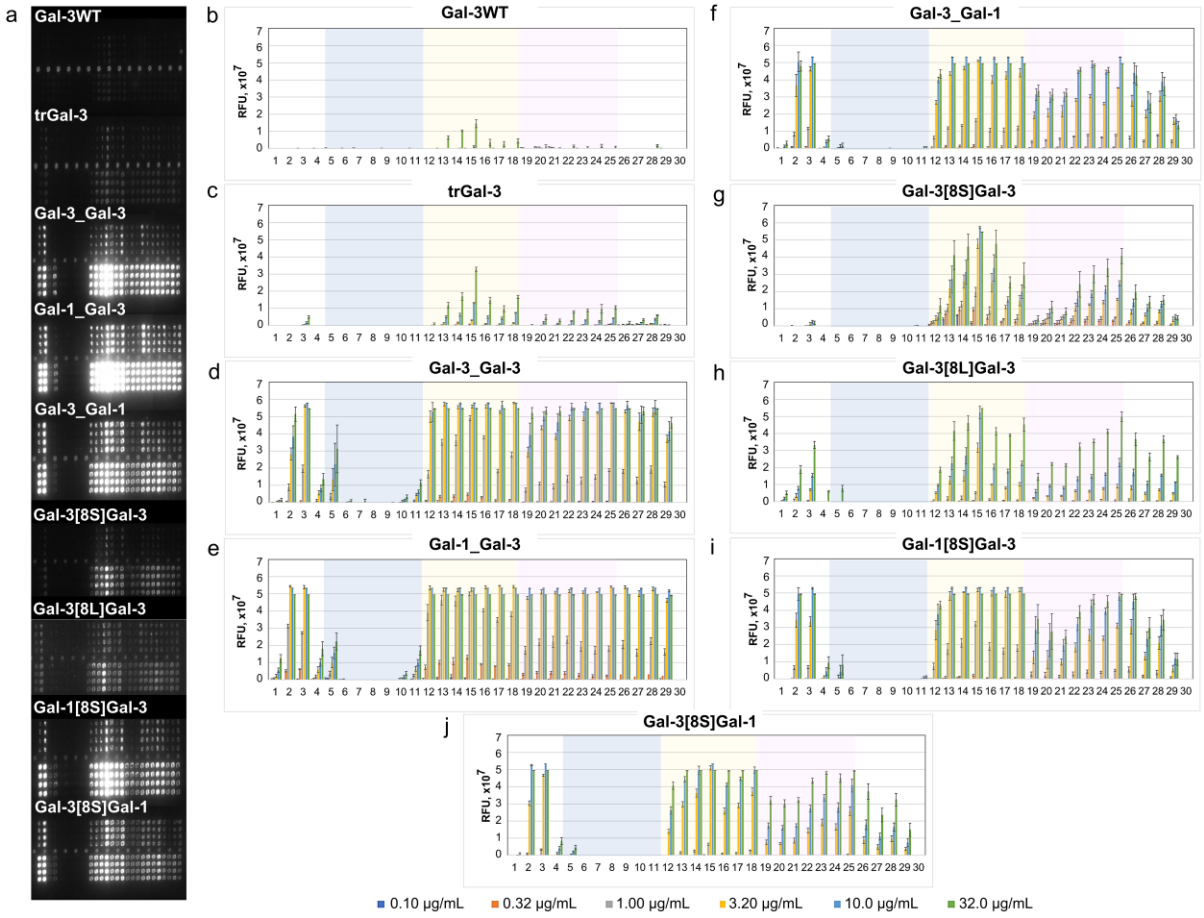

Figure S4. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with 10.0 µg/mL galectin solution (**a**) and relative interaction of **200 µM** core M1 α-DG glycoconjugates with 0.10 µg/mL to 32.0 µg/mL Gal-3WT (**b**), trGal-3 (**c**), Gal-3_Gal-3 (**d**), Gal-1_Gal-3 (**e**), Gal-3_Gal-1 (**f**), Gal-3[8S]Gal-3 (**g**), Gal-3[8L]Gal-3 (**h**), Gal-1[8S]Gal-3 (**i**), and Gal-3[8S]Gal-1 (**j**).

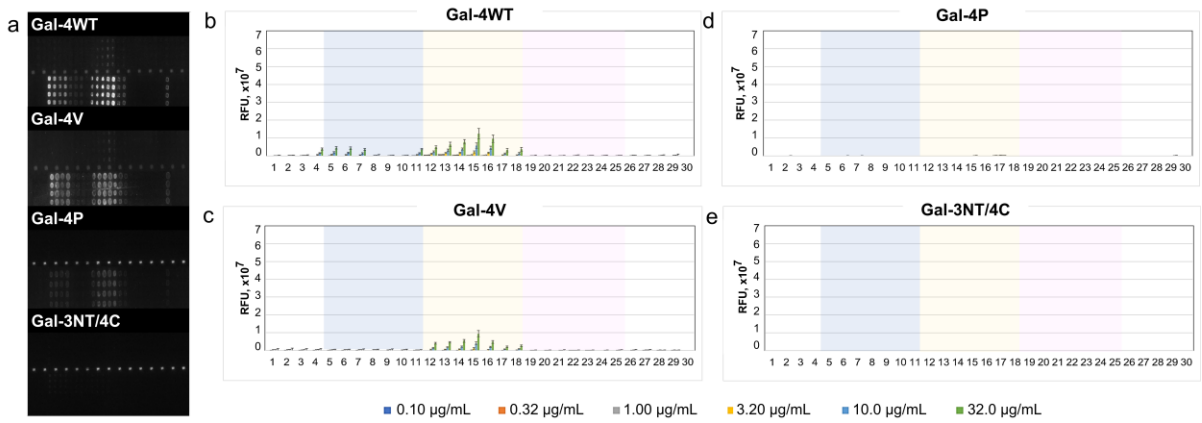

Figure S5. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with 10.0 µg/mL galectin solution (**a**) and relative interaction of **20 µM** core M1 α-DG glycoconjugates with 0.10 µg/mL to 32.0 µg/mL Gal-4WT (**b**), Gal-4V (**c**), Gal-4P (**d**), and Gal-3NT/4C (**e**).

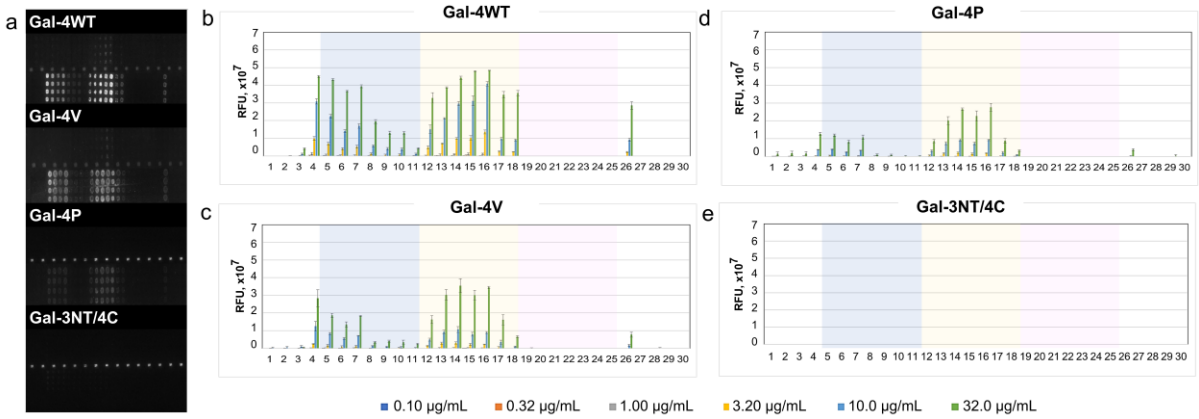

Figure S6. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with 10.0 µg/mL galectin solution (**a**) and relative interaction of **200 µM** core M1 α-DG glycoconjugates with 0.10 µg/mL to 32.0 µg/mL Gal-4WT (**b**), Gal-4V (**c**), Gal-4P (**d**), and Gal-3NT/4C (**e**).

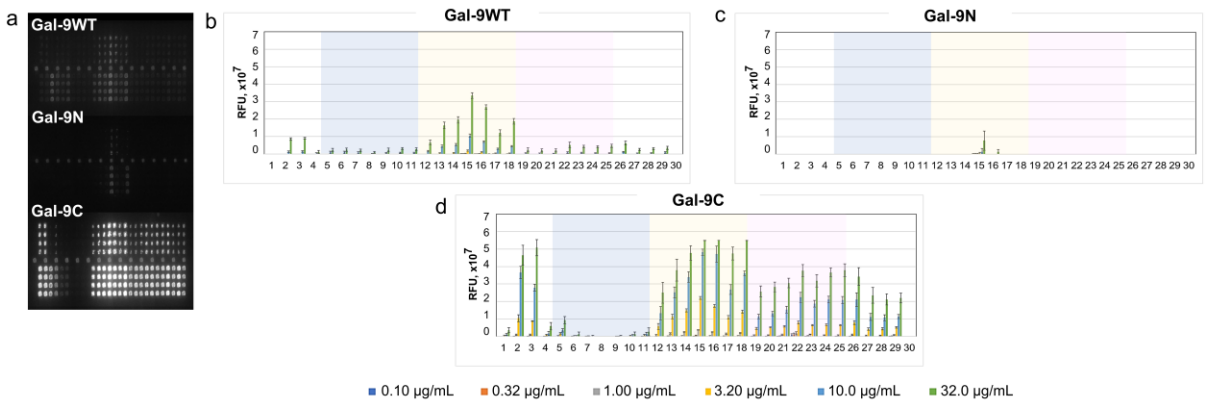

Figure S7. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with 10.0 µg/mL galectin solution (**a**) and relative interaction of **20 µM** core M1 α-DG glycoconjugates with 0.10 µg/mL to 32.0 µg/mL Gal-9WT (**b**), Gal-9N (**c**), and Gal-9C (**d**).

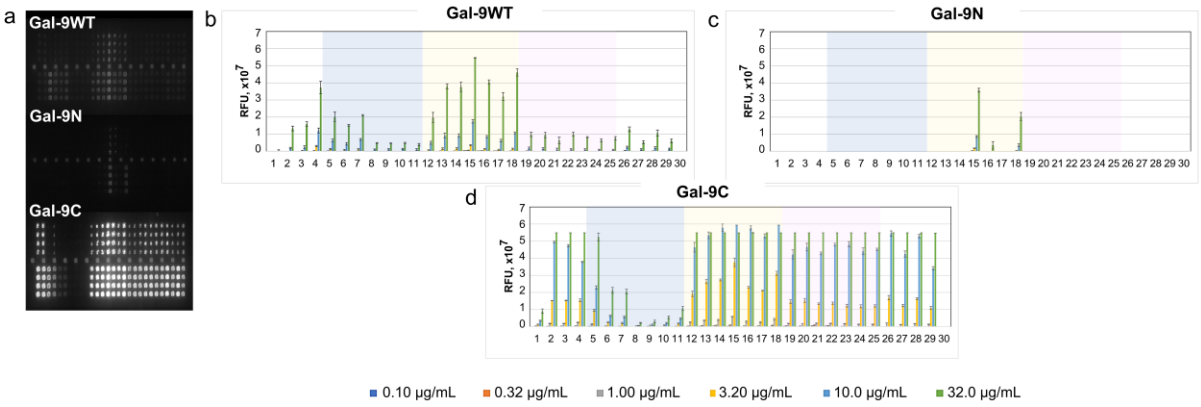

Figure S8. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after treatment with 10.0 µg/mL galectin solution (**a**) and relative interaction of **200 µM** core M1 α-DG glycoconjugates with 0.10 µg/mL to 32.0 µg/mL Gal-9WT (**b**), Gal-9N (**c**), and Gal-9C (**d**).

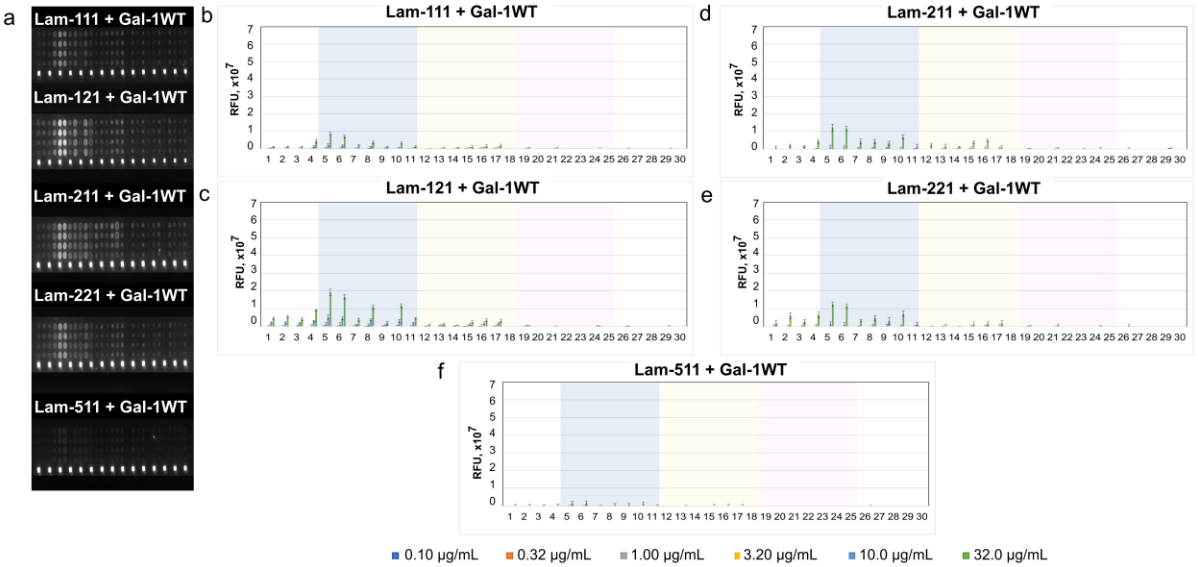

Figure S9. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after 30 min treatment with 32.0 µg/mL **laminin–Gal-1WT** solution (**A**) and relative binding properties of 200 µM core M1 of αDG with 0.10 to 32.0 µg/mL Laminin-111 and $(Gal\text{-}1)_2$[8S] (**B**), Laminin-121 and Gal-1 (**C**), Laminin-211 and Gal-1 (**D**), Laminin-221 and Gal-1 (**E**), Laminin-511 and Gal-1 (**F**).

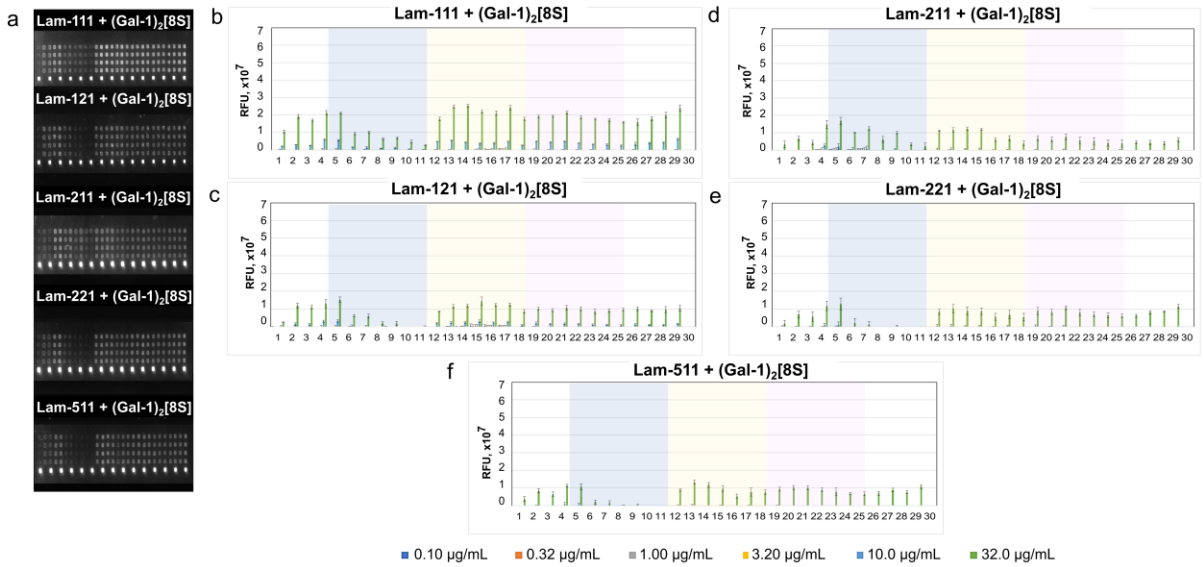

Figure S10. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after 30 min treatment with 32.0 μg/mL **laminin–(Gal-1)$_2$[8S]** solution (**A**) and relative binding properties of 200 μM core M1 of αDG with 0.10 to 32.0 μg/mL Laminin-111 and (Gal-1)$_2$[8S] (**B**), Laminin-121 and Gal-1 (**C**), Laminin-211 and Gal-1 (**D**), Laminin-221 and Gal-1 (**E**), Laminin-511 and Gal-1 (**F**).

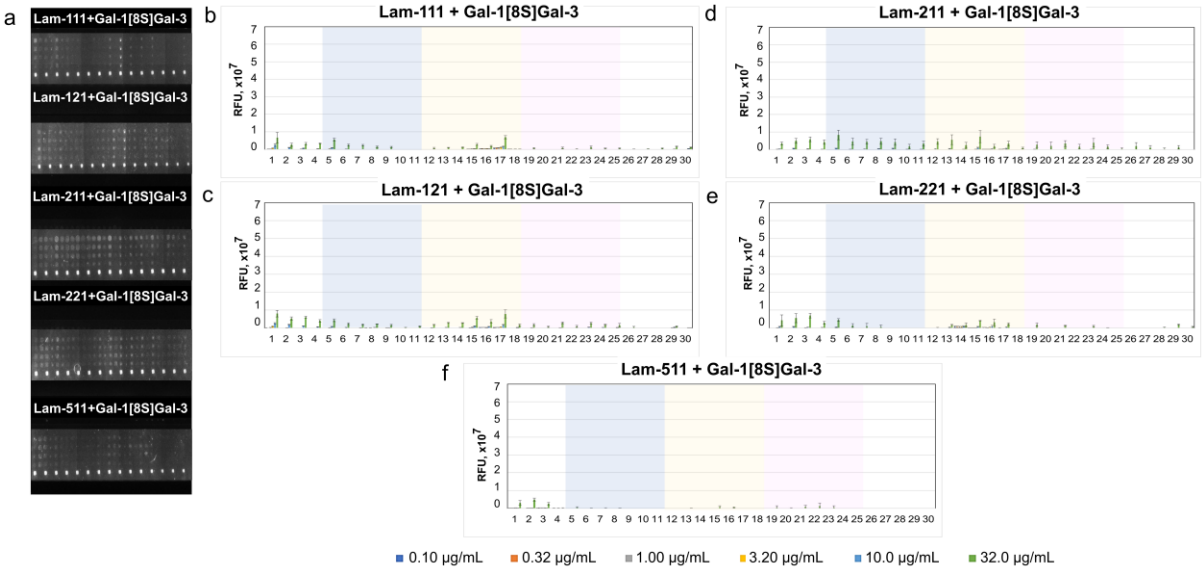

Figure S11. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after 30 min treatment with 32.0 µg/mL **laminin–Gal-1[8S]Gal-3** solution (**A**) and relative binding properties of 200 µM core M1 of αDG with 0.10 to 32.0 µg/mL Laminin-111 and $(Gal\text{-}1)_2$[8S] (**B**), Laminin-121 and Gal-1 (**C**), Laminin-211 and Gal-1 (**D**), Laminin-221 and Gal-1 (**E**), Laminin-511 and Gal-1 (**F**).

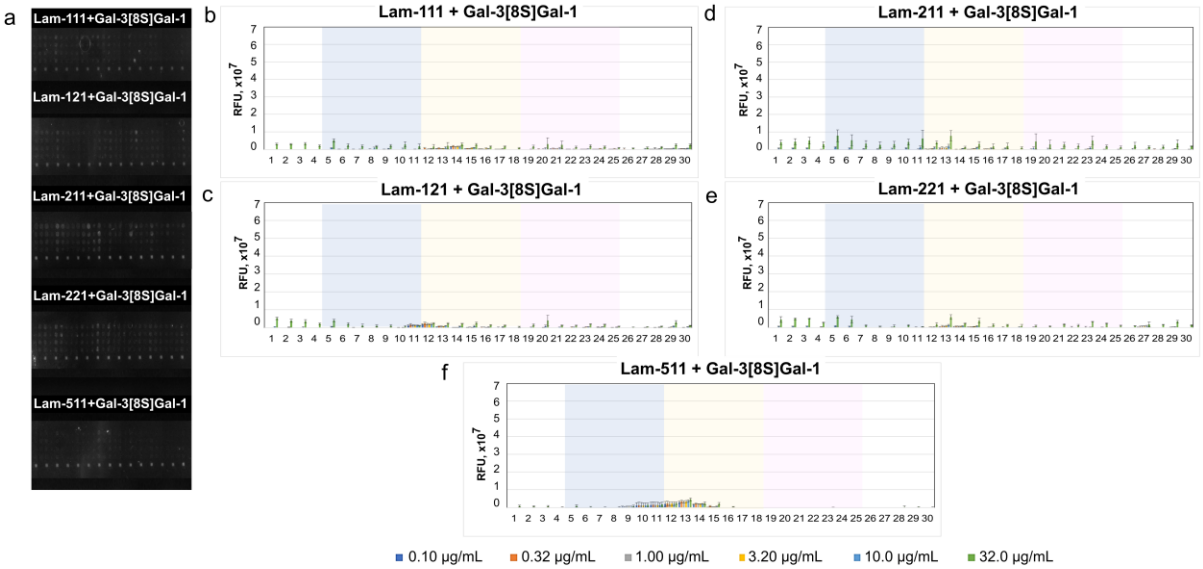

Figure S12. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after 30 min treatment with 32.0 µg/mL **laminin–Gal-3[8S]Gal-1** solution (**A**) and relative binding properties of 200 µM core M1 of αDG with 0.10 to 32.0 µg/mL Laminin-111 and $(Gal\text{-}1)_2[8S]$ (**B**), Laminin-121 and Gal-1 (**C**), Laminin-211 and Gal-1 (**D**), Laminin-221 and Gal-1 (**E**), Laminin-511 and Gal-1 (**F**).

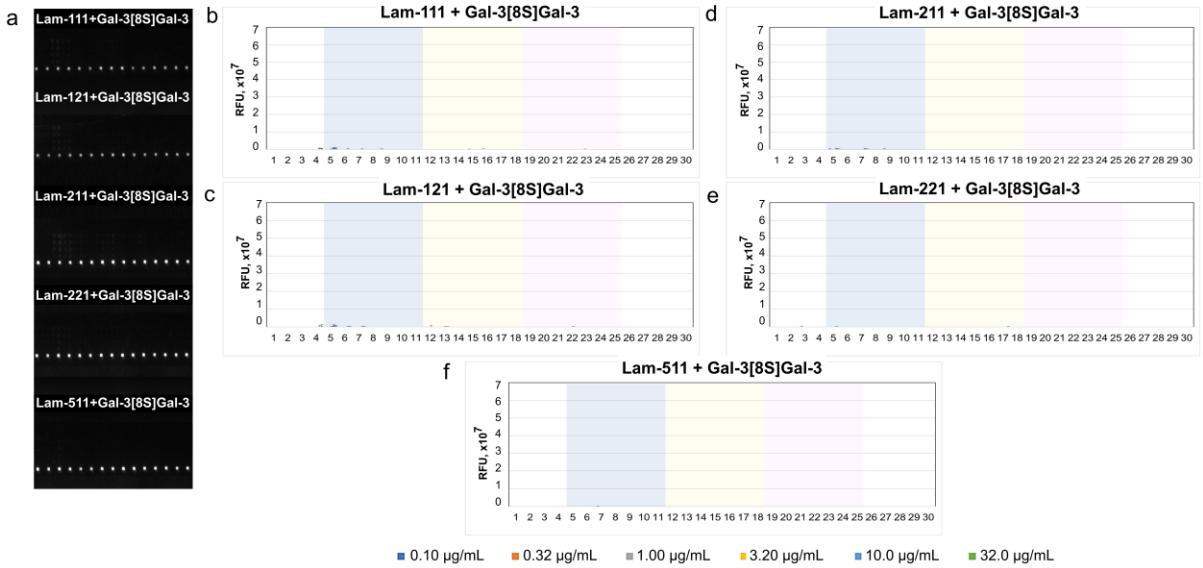

Figure S13. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after 30 min treatment with 32.0 µg/mL **laminin–Gal-3[8S]Gal-3** solution (**A**) and relative binding properties of 200 µM core M1 of αDG with 0.10 to 32.0 µg/mL Laminin-111 and $(Gal\text{-}1)_2[8S]$ (**B**), Laminin-121 and Gal-1 (**C**), Laminin-211 and Gal-1 (**D**), Laminin-221 and Gal-1 (**E**), Laminin-511 and Gal-1 (**F**).

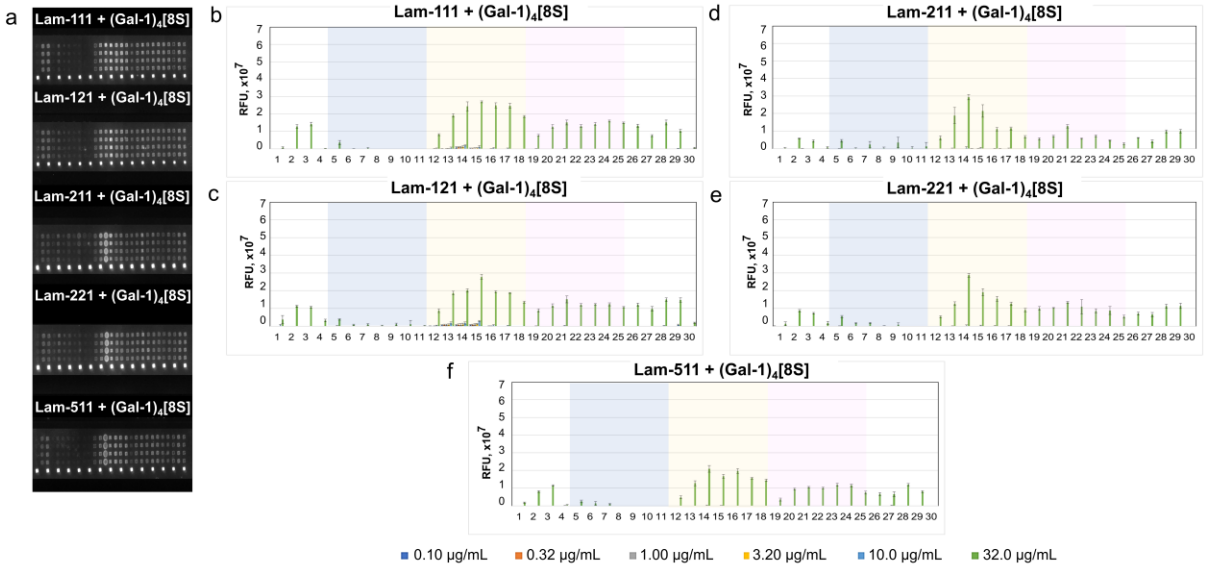

Figure S14. Fluorescence image of core M1 α-DG glycopeptide printed microarray chip taken after 30 min treatment with 32.0 µg/mL **laminin–(Gal-1)$_4$[8S]** solution (**A**) and relative binding properties of 200 µM core M1 of αDG with 0.10 to 32.0 µg/mL Laminin-111 and (Gal-1)$_4$[8S] (**B**), Laminin-121 and Gal-1 (**C**), Laminin-211 and Gal-1 (**D**), Laminin-221 and Gal-1 (**E**), Laminin-511 and Gal-1 (**F**).

5.5. References

1. C. Reily, T. J. Stewart, M. B. Renfrow, J. Novak, Glycosylation in health and disease. *Nat. Rev. Nephrol.* **15**, 346–366 (2019).

2. H. Kaltner, *et al.*, Galectins: their network and roles in immunity/tumor growth control. *Histochem. Cell Biol.* **147**, 239–256 (2017).

3. K. Kasai, Galectins: Quadruple-faced Proteins. *Trends Glycosci. Glycotechnol.* **30**, SE221–SE223 (2018).

4. G. García Caballero, *et al.*, How galectins have become multifunctional proteins. *Histol. Histopathol.* **35**, 509—539 (2020).

5. H.-J. Gabius, J. Roth, An introduction to the sugar code. *Histochem. Cell Biol.* **147**, 111–117 (2017).

6. H.-J. Gabius, S. André, J. Jiménez-Barbero, A. Romero, D. Solís, From lectin structure to functional glycomics: principles of the sugar code. *Trends Biochem. Sci.* **36**, 298–313 (2011).

7. J. C. Manning, *et al.*, Lectins: a primer for histochemists and cell biologists. *Histochem. Cell Biol.* **147**, 199–222 (2017).

8. R. Heusschen, I. A. Schulkens, J. van Beijnum, A. W. Griffioen, V. L. Thijssen, Endothelial LGALS9 splice variant expression in endothelial cell biology and angiogenesis. *Biochim. Biophys. Acta* **1842**, 284–292 (2014).

9. N. Nishi, A. Itoh, H. Shoji, H. Miyanaka, T. Nakamura, Galectin-8 and galectin-9 are novel substrates for thrombin. *Glycobiology* **16**, 15C-20C (2006).

10. G. A. Rabinovich, M. A. Toscano, S. S. Jackson, G. R. Vasta, Functions of cell surface

galectin-glycoprotein lattices. *Curr. Opin. Struct. Biol.* **17**, 513–520 (2007).

11. H. Kaltner, *et al.*, Revealing biomedically relevant cell and lectin type-dependent structure–activity profiles for glycoclusters by using tissue sections as an assay platform. *RSC Adv.* **8**, 28716–28735 (2018).

12. J. Kopitz, *et al.*, Reaction of a Programmable Glycan Presentation of Glycodendrimersomes and Cells with Engineered Human Lectins To Show the Sugar Functionality of the Cell Surface. *Angew. Chemie Int. Ed.* **56**, 14677–14681 (2017).

13. C. Zhou, H. L. Reesink, D. A. Putnam, Selective and Tunable Galectin Binding of Glycopolymers Synthesized by a Generalizable Conjugation Method. *Biomacromolecules* **20**, 3704–3712 (2019).

14. G. Artigas, H. Hinou, F. Garcia-Martin, H.-J. Gabius, S.-I. Nishimura, Synthetic Mucin-Like Glycopeptides as Versatile Tools to Measure Effects of Glycan Structure/Density/Position on the Interaction with Adhesion/Growth-Regulatory Galectins in Arrays. *Chem. Asian J.* **12**, 159–167 (2017).

15. N. Nishi, *et al.*, Functional and structural bases of a cysteine-less mutant as a long-lasting substitute for galectin-1. *Glycobiology* **18**, 1065–1073 (2008).

16. P. Bättig, P. Saudan, T. Gunde, M. F. Bachmann, Enhanced apoptotic activity of a structurally optimized form of galectin-1. *Mol. Immunol.* **41**, 9–18 (2004).

17. L. A. Earl, S. Bi, L. G. Baum, Galectin multimerization and lattice formation are regulated by linker region structure. *Glycobiology* **21**, 6–12 (2011).

18. S. Tasumi, G. R. Vasta, A galectin of unique domain organization from hemocytes of the Eastern oyster (Crassostrea virginica) is a receptor for the protistan parasite

Perkinsus marinus. *J. Immunol.* **179**, 3086–3098 (2007).

19. X. Song, *et al.*, An immune responsive multidomain galectin from bay scallop Argopectens irradians. *Fish Shellfish Immunol.* **28**, 326–332 (2010).

20. M. M. Fettis, S. A. Farhadi, G. A. Hudalla, A chimeric, multivalent assembly of galectin-1 and galectin-3 with enhanced extracellular activity. *Biomater. Sci.* **7**, 1852–1862 (2019).

21. S. A. Farhadi, M. M. Fettis, R. Liu, G. A. Hudalla, A Synthetic Tetramer of Galectin-1 and Galectin-3 Amplifies Pro-apoptotic Signaling by Integrating the Activity of Both Galectins . *Front. Chem.* **7** (2020).

22. J. M. Ervasti, K. Ohlendieck, S. D. Kahl, M. G. Gaver, K. P. Campbell, Deficiency of a glycoprotein component of the dystrophin complex in dystrophic muscle. *Nature* **345**, 315–319 (1990).

23. T. ENDO, Mammalian *O*-mannosyl glycans: Biochemistry and glycopathology. *Proc. Japan Acad. Ser. B* **95**, 39–51 (2019).

24. T. Yoshida-Moriguchi, K. P. Campbell, Matriglycan: a novel polysaccharide that links dystroglycan to the basement membrane. *Glycobiology* **25**, 702–713 (2015).

25. P. M. Van Ry, *et al.*, ECM-Related Myopathies and Muscular Dystrophies: Pros and Cons of Protein Therapies. *Compr. Physiol.* **7**, 1519–1536 (2017).

26. L. L. Villones, *et al.*, Exploring the In situ pairing of human galectins toward synthetic O-mannosylated core M1 glycopeptides of α-dystroglycan. *Sci. Rep.* **12**, 17800 (2022).

27. A.-K. Ludwig, H. Kaltner, J. Kopitz, H.-J. Gabius, Lectinology 4.0: Altering modular (ga)lectin display for functional analysis and biomedical applications. *Biochim.*

Biophys. Acta - Gen. Subj. **1863**, 935–940 (2019).

28. A.-K. Ludwig, *et al.*, Design–functionality relationships for adhesion/growth-regulatory galectins. *Proc. Natl. Acad. Sci.* **116**, 2837 LP – 2842 (2019).

29. T. J. Kutzner, *et al.*, How altering the modular architecture affects aspects of lectin activity: case study on human galectin-1. *Glycobiology* **29**, 593–607 (2019).

30. G. García Caballero, *et al.*, Influence of protein (human galectin-3) design on aspects of lectin activity. *Histochem. Cell Biol.* **154**, 135–153 (2020).

31. S. André, G.-N. Wang, H.-J. Gabius, P. V Murphy, Combining glycocluster synthesis with protein engineering: an approach to probe into the significance of linker length in a tandem-repeat-type lectin (galectin-4). *Carbohydr. Res.* **389**, 25–38 (2014).

32. D. Solís, *et al.*, N-domain of human adhesion/growth-regulatory galectin-9: preference for distinct conformers and non-sialylated N-glycans and detection of ligand-induced structural changes in crystal and solution. *Int. J. Biochem. Cell Biol.* **42**, 1019–1029 (2010).

33. Q. Xiao, *et al.*, Exploring functional pairing between surface glycoconjugates and human galectins using programmable glycodendrimersomes. *Proc. Natl. Acad. Sci.* **115**, E2509 LP-E2518 (2018).

34. H. Hinou, *et al.*, Synthetic glycopeptides reveal specific binding pattern and conformational change at O-mannosylated position of α-dystroglycan by POMGnT1 catalyzed GlcNAc modification. *Bioorg. Med. Chem.* **27**, 2822–2831 (2019).

35. T. Matsushita, *et al.*, A straightforward protocol for the preparation of high performance microarray displaying synthetic MUC1 glycopeptides. *Biochim. Biophys.*

Acta - Gen. Subj. **1840**, 1105–1116 (2014).

36. Y. Nonaka, *et al.*, Crystal structure and conformational stability of a galectin-1 tandem-repeat mutant with a short linker. *Glycobiology* **32**, 251–259 (2022).

37. S. Bi, L. A. Earl, L. Jacobs, L. G. Baum, Structural features of galectin-9 and galectin-1 that determine distinct T cell death pathways. *J. Biol. Chem.* **283**, 12248–12258 (2008).

38. S. Vértesy, *et al.*, Structural significance of galectin design: impairment of homodimer stability by linker insertion and partial reversion by ligand presence. *Protein Eng. Des. Sel.* **28**, 199–210 (2015).

39. N. Ahmad, *et al.*, Galectin-3 Precipitates as a Pentamer with Synthetic Multivalent Carbohydrates and Forms Heterogeneous Cross-linked Complexes*. *J. Biol. Chem.* **279**, 10841–10847 (2004).

40. Y.-H. Lin, *et al.*, The intrinsically disordered N-terminal domain of galectin-3 dynamically mediates multisite self-association of the protein through fuzzy interactions. *J. Biol. Chem.* **292**, 17845–17856 (2017).

41. J. Seetharaman, *et al.*, X-ray Crystal Structure of the Human Galectin-3 Carbohydrate Recognition Domain at 2.1-Å Resolution*. *J. Biol. Chem.* **273**, 13047–13052 (1998).

42. P. M. Collins, K. Bum-Erdene, X. Yu, H. Blanchard, Galectin-3 interactions with glycosphingolipids. *J. Mol. Biol.* **426**, 1439–1451 (2014).

43. M. J. Moure, *et al.*, Selective 13C-Labels on Repeating Glycan Oligomers to Reveal Protein Binding Epitopes through NMR: Polylactosamine Binding to Galectins. *Angew. Chemie Int. Ed.* **60**, 18777–18782 (2021).

44. S. Kamitori, Three-Dimensional Structures of Galectins. *Trends Glycosci. Glycotechnol.* **30**, SE41–SE50 (2018).

45. P. M. Van Ry, R. D. Wuebbles, M. Key, D. J. Burkin, Galectin-1 Protein Therapy Prevents Pathology and Improves Muscle Function in the mdx Mouse Model of Duchenne Muscular Dystrophy. *Mol. Ther.* **23**, 1285–1297 (2015).

46. R. D. Wuebbles, *et al.*, Human Galectin-1 Improves Sarcolemma Stability and Muscle Vascularization in the mdx Mouse Model of Duchenne Muscular Dystrophy. *Mol. Ther. Methods Clin. Dev.* **13**, 145–153 (2019).

47. T. L. M. Thurston, M. P. Wandel, N. von Muhlinen, Á. Foeglein, F. Randow, Galectin 8 targets damaged vesicles for autophagy to defend cells against bacterial invasion. *Nature* **482**, 414–418 (2012).

48. B.-W. Kim, S. Beom Hong, J. Hoe Kim, D. Hoon Kwon, H. Kyu Song, Structural basis for recognition of autophagic receptor NDP52 by the sugar receptor galectin-8. *Nat. Commun.* **4**, 1613 (2013).

49. F. A. van den Brûle, C. Buicu, M. E. Sobel, F. T. Liu, V. Castronovo, Galectin-3, a laminin binding protein, fails to modulate adhesion of human melanoma cells to laminin. *Neoplasma* **42**, 215–219 (1995).

50. I. Kuwabara, F. T. Liu, Galectin-3 promotes adhesion of human neutrophils to laminin. *J. Immunol.* **156**, 3939–3944 (1996).

51. J. Kopitz, S. Ballikaya, S. André, H.-J. Gabius, Ganglioside GM1/Galectin-Dependent Growth Regulation in Human Neuroblastoma Cells: Special Properties of Bivalent Galectin-4 and Significance of Linker Length for Ligand Selection. *Neurochem. Res.* **37**, 1267–1276 (2012).

52. A. Kuno, *et al.*, Evanescent-field fluorescence-assisted lectin microarray: a new strategy for glycan profiling. *Nat. Methods* **2**, 851–856 (2005).

Chapter 6.

Summary and Prospects

Aberrant O-mannosylation in mammalian cells causes multiple forms of muscular and neurological disorders, indicating that this unique form of glycosylation has an essential, yet less understood, role in human health and disease. Combining the availability of the array platform with a core M1-type glycopeptide library and a panel of fluorescent galectins, the binding was detected and the profiles were mapped. The type of glycan, position, and density along the α-DG peptide scaffold, as well as the galectin architecture, determined this binding event. Human galectins can recognize *O*-Man LacNAc-terminated glycoconjugates, making their respective in situ contacts possible. The presence of an α2,3-sialylated terminus led to a major reduction in the affinity of galectin, suggesting that this type of extension can fine-tune or act as an on/off switch of galectin activity towards *O*-Man glycans. These interactions were significantly inhibited by lactose, establishing that the α-DG core M1-type glycans bind to the canonical sugar-binding site (S-face) of galectin, thus serving as a receptor for galectins. The interaction with the peptide region of α-DG, which was not entirely inhibited by lactose, strongly implies that this peptide-galectin interaction occurs via specific hydrophobic interactions that are not dependent on the S-face of galectins. Furthermore, based on our microarray findings, the 1H-15N HSQC NMR data showed that LacNAc-terminated core M1 glycopeptide could interact with the S-face (carbohydrate-protein) and F-face (peptide-protein) of Gal-1, which was diminished by the α2,3-sialylation of this glycoconjugate. Moreover, Gal-1 effectively *trans*-bridge various laminins (111, 121, 211, and 221) to the α-DG peptide and core M1 glycopeptides in the array. This Gal-1 *trans*-bridging activity was completely abrogated by the α2,3-sia extension of the LacNAc core M1 glycopeptides, indicating that Gal-1 binding to the polyLacNAc side chain of laminin is preferred over α-DG core M1 sialyl-LacNAc ligands.

After demonstrating that core M1 α-DG core glycoconjugates are receptors for galectins, the effect of alteration of galectin design and valency on the binding event was

evaluated utilizing rational protein engineering and microarray techniques. These experimental setups demonstrated that the alteration of the galectin structures can give additional insights into the preferential modular architecture and binding behavior of this lectin family towards the prepared ligands. In addition, rational protein engineering was a useful tool in redesigning lectins with possibly higher therapeutic potentials than their wild-type counterpart. Here, Gal-1 has proof-of-principle character, where we were able to enhance its *cis*-binding and *trans*-bridging activities with the core M1 α-DG core glycocopeptides compared to its parent protein.

However, unlike cellular or in vivo assays, our experimental setup in this study did not provide biosignals for the observed cis/trans-bridging activities of galectin with α-DG glycopeptides. The extension of core M3 in α-DG is initiated by FKTN to form the matriglycan [-GlcAβ1,3Xylα1,3-]n that serves as a ligand for laminin linking DG to the basement membrane. Mutation in FKTN gene results in FCMD, one of the most common forms of muscular dystrophy in Japan. Since the matriglycan structure is completely abrogated, the core M1 structures and the α-DG peptide could serve as contact points of Gal-1 and its variants to *trans*-bridge with laminin reestablishing the connection of DG and basement membrane. This inspires further analysis of the (patho)physiological functions of galectin–core M1 α-DG interactions, specifically in the muscle and neural cells.

Acknowledgements

Primarily, I would like to thank the Almighty Father, the source of my strength, courage, knowledge and wisdom.

To my family thank you very much for the love and support you have poured unto me. No words can express how grateful I am to be part of this family.

I offer my deepest appreciation to my thesis supervisor, Prof. Hiroshi Hinou for the encouragement and enthusiasm which allowed me to finish this project. I will not be able to finish this research without your kind and patient instructions.

I would also like to express my gratitude to our collaborators; the late Prof. Hans-Joachim Gabius, Prof. Shin-Ichiro Nishimura, Prof. Tomoyasu Aizawa, Dr. Anna-Kristin Ludwig, Dr. Hiroyuki Kumeta, Mr. Seiya Kikuchi, and Dr. Rika Ochi for sharing their expertise on conducting the experiments needed to complete this study. It was a great privilege and honor to work and study under their guidance.

I am also extremely thankful to my panel members, Prof. Shin-Ichiro Nishimura, Prof. Tomoyasu Aizawa, and Prof. Ryota Uehara, for sparing their valuable time in criticizing constructively my research and in reviewing and polishing the manuscript which improved the quality of this research.

To all my friends, members of the Laboratory of Advanced Chemical Biology, and many people who have been a part of my graduate education here in Hokkaido University, I am truly grateful to all of you guys.

Lastly, to Hokkaido University, the International Graduate Program, and the Japanese Ministry of Education, Culture, Sports, Science, and Technology scholarship program for the financial support throughout my graduate studies and research funding which made this study possible.