167

Chapter 7

Carbohydrate Microarrays for Lectin Characterization and Glyco-Epitope Identification

Denong Wang[a] and Albert M. Wu[b]

[a]*Carbohydrate Microarray Laboratory, Department of Genetics, and Neurology and Neurological Sciences, Stanford University School of Medicine, Beckman Center B007, Stanford, CA 94305-5318, USA*
[b]*Glyco-Immunochemistry Research Laboratory, Institute of Molecular and Cellular Biology, College of Medicine, Chang-Gung University, Kwei-san, Tao-yuan, 333, Taiwan*

1. Introduction

Lectins are an important class of proteins or glycoproteins of non-immune origin that bind non-covalently to characteristic carbohydrate structures with specificity or selectivity. There are many ways to classify lectins. Given the fact that a wide range of living organisms, from microbes to mammals, produce lectins, they can be divided according to species origin, such as microbial lectins, plant lectins, invertebrate lectins and vertebrate lectins. Based on the similarities in sequence of homology and activity, lectins can also be grouped into subfamilies. There are at least five groups of vertebrate lectins, such as C-type (calcium dependent), P-type, S-type, I-type, and pentraxins [1]. Cellular lectin molecules or receptors with lectin-like carbohydrate-binding domains play important roles in cell signaling, protein intracellular trafficking, and cell–cell communication [2–5].

Lectins can also be classified according to their binding specificity or selectivity with carbohydrates. This type of classification is helpful for selection of lectins as structural probes in biomedical applications. By the early seventies, the carbohydrate-binding activity of a lectin was commonly described in terms of monosaccharide specificity. Experimentally, a type of inhibition assay, such as agglutination inhibition, was performed to identify the monosaccharide that most effectively inhibited the binding reaction. Many lectins are able to cross-react with a panel of sugar chains with a common terminal sugar residue. Lately, disaccharides are considered to be a better way to define a lectin's specificity.

However, some lectin-binding sites accommodate glyco-epitopes larger than mono- and disaccharides and recognize conformational epitopes. The latter may consist of much longer oligosaccharide or polysaccharide chains. There is also evidence that lectins recognize certain cluster configurations of complex carbohydrates with high selectivity and avidity.

X-ray crystallographic studies identified some common characteristics in the binding pockets of lectins and antibodies [2, 6, 7]. Aromatic amino acid residues appear frequently in carbohydrate-binding sites of various fine specificities. Such residues are large and participate in a wide variety of van der Waals and electrostatic interactions. The majority of crystal structures so far resolved are associated with a network of H-bonds, frequently with water molecules, in their combining sites. Such bound water molecules play important roles at the contacting interface between carbohydrate ligands and receptors. In addition, polyamphiphilic surfaces were identified in the interfaces of a number of carbohydrate ligands and binding sites. Presence of relatively rigid structures is required to produce an amphiphilic surface in solution. Carbohydrate molecules with branched termini, such as many blood group substances, which are formed by oligosaccharide chains, are favorable for the generation of such contact surfaces.

Much remains to be learned regarding the specificity and cross-reactivity of the carbohydrate–lectin interactions and their functional outcome in biological systems. Challenging issues in lectin characterization may include but are not limited to (a) the spectrum or repertoire of carbohydrate structures that are reactive with given lectins, (b) conformational properties of glyco-epitopes that are selectively recognized by lectins, and (c) the cluster effect in lectin–carbohydrate interactions, especially those that take place on cell surfaces and trigger the events of intracellular cell signaling, cell differentiation and proliferation. Recent development of microarray-based broad-range binding assays offers new tools to facilitate these investigations.

In this chapter, we summarize the concept and the updated information regarding the specificity-based lectin classification (Section 2). Then, we introduce a practical platform of carbohydrate microarrays that is likely useful for lectin characterization and classification (Section 3). Lastly, we discuss a few examples to illustrate the application of this technology in lectin-related experimental investigations (Section 4).

2. The Specificity-Based Lectin Classification

Lectins are functionally classified based on their relative binding reactivities with the structural units of carbohydrate or glyco-epitopes. They are grouped according to their monosaccharide specificities and then further sub-grouped based on their reactivities with more complex structures. As listed below, carbohydrate specificities of biomedically important lectins are classified into six

groups according to their specificities to monosaccharides. They are further sub-grouped based on the binding affinities to (a) GalNAcα1→*O* to Ser(Thr) of the peptide chain; (b) disaccharides; (c) trisaccharides; and (d) the number and the location of LFucα1→linked to oligosaccharides. These structures are frequently found in soluble glycoproteins and as cell surface glycoconjugates in mammals [8–11]. This scheme of lectin classification is, thus, practically useful in bio-medical application of lectins.

2.1. Group I. GalNAc-specific agglutinins

(1) **F/A**, GalNAcα1→3GalNAc (**F**orssman) and GalNAcα1→3Gal (blood group **A** determinant disaccharide) specific agglutinins – *Dolichos biflorus* (DBA) and *Helix pomatia* (HPA).
(2) **A**, GalNAcα1→3Gal (blood group **A** determinant disaccharide) – specific agglutinins – Soybean (SBA) and Lima bean (LBA).
(3) **Tn**, GalNAcα1→Ser/Thr-specific agglutinins – *Vicia villosa* B_4 (VVL-B_4) and *Salvia sclarea* (SSA).

2.2. Group II. Gal-specific agglutinins

(1) **T**, Galβ1→3GalNAc (**T**$_α$, Galβ1→3GalNAcα1→ the mucin-type sugar sequence on human erythrocyte membrane or **T**$_β$, Galβ1→3GalNAcβ1→ at the terminal non-reducing end of the gangliosides) specific agglutinins – Peanut (PNA), and *Bauhinia purpurea alba* (BPA).
(2) **I/II**, Galβ1→3(4)GlcNAcβ1→(Lacto-*N*-biose/*N*-acetyllactosamine) spe-cific agglutinins (Human blood group type **I** (Galβ1→3GlcNAc) and type **II** (Galβ1→4GlcNAc) carbohydrate sequences [16–18], the disaccharide residues at the non-reducing end of the carbohydrate chains derived from either *N*-glycosidic or *O*-glycosidic linkages.) – *Ricinus communis* agglu-tinin (RCA$_1$), *Datura stramonium* (thorn apple, TAA) and wheat germ agglutinin (WGA).
(3) **B**, Galα1→3Gal (Human blood group **B** disaccharide-*Griffonia (Bandeiraea) simplicifolia* B_4 (GSI-B_4).
(4) **E**, Galα1→4Gal, blood group P^k active disaccharide – Abrin-a and Mistletoe lectin-I (ML-I).

2.3. Group III. Man and/or Glc-specific agglutinins

Man-linked oligosaccharide-specific agglutinins – *Concanavalin ensiformis* (Jack bean, ConA) and *Lens culinaris* (LCA).

2.4. Group IV. GlcNAc, and/or Galβ1→4GlcNAcβ1→linked specific agglutinins

Chitin oligosaccharide-specific agglutinins – WGA and *Griffonia (Bandeiraea) simplicifolia* II (GSA-II).

2.5. Group V. LFuc-specific agglutinins

(1) Monofucosyl-specific agglutinins – *Ulex europaeus* I (UEA-I) and UEA-II.
(2) Difucosyl-specific agglutinins – *Griffonia (Bandeiraea) simplicifolia* IV (GSA-IV).
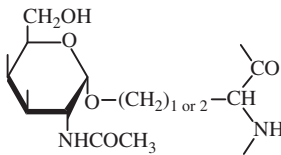
2.6. Group VI. Sialic acid specific agglutinins

2.6.1. Limulus polyphemus agglutinin (LPA).

Glycoproteins and glycosphingolipids contain many carbohydrate epitopes or crypto-glyco-epitopes of biomedical importance. They are present on the cell surface and function as receptors in various life processes. Many known glyco-epitopes exist in soluble or gel form and serve as biological lubricants or as barriers against microbial invasion. During the past decade, 11 mammalian structural units have been used to express the binding domain of Gal- and GalNAc-specific lectins. They are **F**, GalNAcα1→3GalNAc; **A**, GalNAcα1→3Gal; **T**, Galβ1→3GalNAc; **I**, Galβ1→3GlcNAc; **II**, Galβ1→4GlcNAc; **B**, Galα1→3Gal; **E**, Galα1→4Gal; **L**, Galβ1→4Glc; **P**, GalNAcβ1→3Gal; **S**, GalNAcβ1→4Gal and **Tn**, GalNAcα1→Ser(Thr) (Fig. 1 and Table 1). Except **L** and **P**, all of the units can be found in glycoproteins. **Tn**, which is an important marker for breast/colon cancer and vaccine development, exists only in *O*-glycans. Natural **Tn** glycoprotein, the simplest mammalian *O*-glycan, is exclusively expressed in the armadillo salivary gland. Antifreeze glycoprotein is composed of repeating units of **T**. *Pneumococcus* type 14 capsular polysaccharide, which has uniform **II** disaccharide as carbohydrate side chains. Asialo human α_1-acid glycoprotein and asialo fetuin provide multi-antennary **II** structures. Human ovarian cyst glycoproteins, which belong to the complex type of glycoform, comprise most of the structural units. To facilitate the selection of lectins that could serve as structural probes, the carbohydrate-binding properties of Gal/GalNAc-reactive lectins have been classified according to their highest affinity for structural units and their binding profiles are expressed in decreasing order of reactivity. The source and the structural relationship of the proposed mammalian structural units (lectin determinants) for Gal- and GalNAc-specific lectins, which are shown in Fig.1 as part of carbohydrate structures, are shown in Fig. 2 and Table 1. Most of the lectin $\mathbf{F}_\beta$ and $\mathbf{T}_\beta$ determinants are found in glycosphingolipids [12–14], and the other determinants are present in mammalian glycoproteins [9, 15–19], especially in

the human blood group A, B, H, Le^a, Le^b, and Ii active glycoproteins prepared from human ovarian cyst fluid [15, 18] and salivary glycoproteins [20]. The reactivities of lectin determinants represent a combined result of the binding of individual sugars. The contribution of each sugar to the binding is not necessarily equal, and is different among lectins, for example, both *Amaranthus caudatus* (ACL) and PNA are Galβ1→3GalNAc specific [8, 21]. However, the inhibitory profile of the monosaccharides with these two lectins is quite different;

I. GalNAcα1→ (Tn, A, F$_\alpha$ and F$_\beta$)

Tn.

F.

Tn, GalNAcα1→ Ser/Thr
(Tn, in *O*-linked glycoprotein)
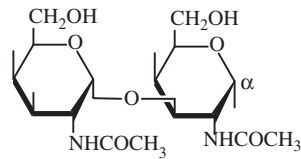
F$_\alpha$, GalNAcα1→3GalNAcα1→
(Core 5, in *O*-linked glycoprotein)

A.

A$_\beta$, GalNAcα1→3Galβ1→
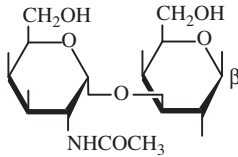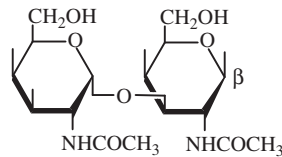
F$_\beta$, GalNAcα1→3GalNAcβ1→
(Terminal disaccharide at the nonreducing end of Forssman glycotope)

Group of GalNAcα1 structural units (Haworth Projection)

II. GalNAcβ1→ (P$_\alpha$ and S$_\beta$)
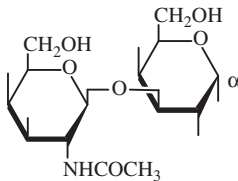
P.

S.

P$_\alpha$, GalNAcβ1→3Galα1→

S$_\beta$, GalNAcβ1→4Galβ1→
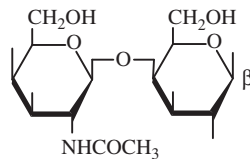
Group of GalNAcβ1→ structural units (Haworth Projection)

Figure 1. Mammalian glycoconjugates structural units used to express and classify the carbohydrate specificity of lectins.

III. Galα1→ (B and E)
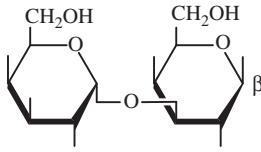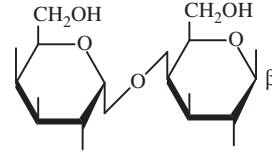
B.   **E.**

B_β, Galα1→3Galβ1→   **E_β**, Galα1→4Galβ1→

Group of Galα1→ structural units (Haworth Projection)

IV. Galβ1→ (Tα, T_β, L, I and II)

T.   **I.**
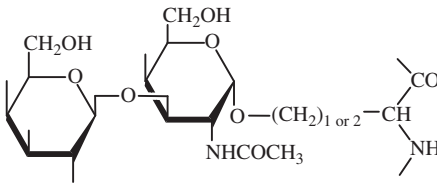
T_α, Galβ1→3GalNAcα1→Ser/Thr   **I_β**, Galβ1→3GlcNAcβ1
(**T**, in *O*-linked glycoprotein)
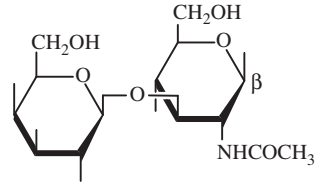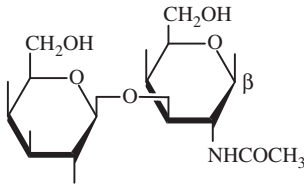
T_β, Galβ1→3GalNAcβ1→   **II_β**, Galβ1→4GlcNAcβ1→
(Terminal disaccharide at the asialo GM1; GSL)
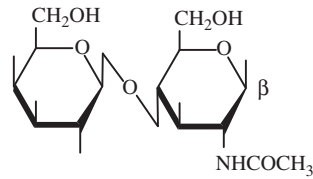
L.

L_β, Galβ1→4Glcβ1→(GSL)
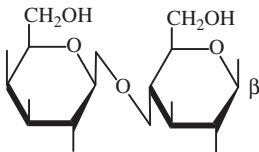
Group of Galβ1→ structural units (Haworth Projection)

Figure 1. (Continued).

Glycosphingolipids

Forssman
pentasaccharide

(F_{penta} or $F_\beta P E_L$)

Asialo GM1

$$F_\beta \qquad E_\beta$$
$$GalNAc\alpha1\rightarrow3GalNAc\beta1\rightarrow3Gal\alpha1\rightarrow4Gal\beta1\rightarrow4Glc\beta1\rightarrow1Cer$$
$$F \qquad P_\alpha \qquad L_\beta$$

$$T_\beta \qquad L_\beta$$
$$Gal\beta1\rightarrow3GalNAc\beta1\rightarrow4Gal\beta1\rightarrow4Glc\beta1\rightarrow1Cer$$
$$S_\beta \quad E_\beta(P^k)$$

Ceramide trihexoside

$$Gal\alpha1\rightarrow4Gal\beta1\rightarrow4Glc\beta1\rightarrow1Cer$$
$$L_\beta$$
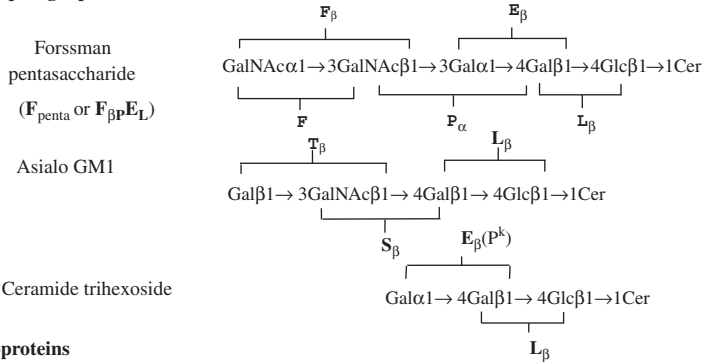
Glycoproteins

a. A, B, A_h, F_α, T_α and Tn structural units of human ovarian cyst (HOC)

$$F_\alpha$$
$$GalNAc\alpha1\rightarrow3GalNAc\alpha1\rightarrow Ser/Thr$$
$$Tn$$

A_h

A   HOC

Fuc
GalNAc$\alpha1\rightarrow3$ $\downarrow\alpha1,2$

Gal$\beta1\rightarrow3(4)$GlcNAc - - - - - - - Gal$\beta1\rightarrow3$GalNAc$\alpha1\rightarrow$Ser/Thr

Gal$\alpha1\rightarrow3$

B   **I (II)**   **Tα**

Non-reducing terminal   Reducing terminal

T   **Tn**

b. Asialo Sd(a^+) THGP

$$S_\beta \qquad II_\beta$$
$$GalNAc\beta1\rightarrow4Gal\beta1\rightarrow4GlcNAc$$
$$\downarrow\beta1,6$$
$$GalNAc\beta1\rightarrow4Gal\beta1\rightarrow4GlcNAc\beta1\rightarrow2Man\alpha1\rightarrow6Man\beta1\rightarrow4GlcNAc\beta1\rightarrow4GlcNAc\beta1\rightarrow Asn$$
$$S_\beta \qquad II_\beta \qquad \uparrow\alpha1,3$$
$$GalNAc\beta1\rightarrow4Gal\beta1\rightarrow4GlcNAc\beta1\rightarrow2Man$$
$$S_\beta \qquad II_\beta \qquad \uparrow\beta1,4$$
$$GalNAc\beta1\rightarrow4Gal\beta1\rightarrow4GlcNAc$$
$$S_\beta \qquad II_\beta$$
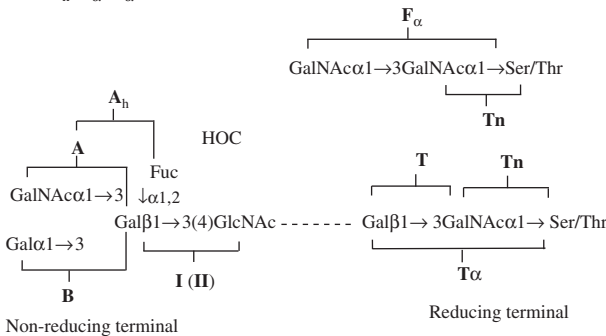
Figure 2. Mammalian structural units (lectin determinants) present in glycoconjugates. Lectin **F** determinant is found in glycosphingolipid (the principal glycolipid of mammalian tissues) of the tissues of the guinea pig, horse, cat, and chicken. It can also be found on the surface of some bacteria, viruses, human gastric carcinoma, and colon tumors [14, 15, 18, 38]. All (**F**, **A/A_h**, **B**, **I/II**, **T**, **Tn**) determinants are found in human blood groups A, B, H, Le^a, Le^b, and Ii active glycoconjugates prepared from human ovarian cyst fluid (HOC) [16, 18, 22, 42]. The Gal $\beta1\rightarrow3$GalNAc group at the non-reducing end of asialo-GM$_1$ is also considered as a lectin T_β determinant. Determinants of lectins **I** and **II** discussed in this article are equivalent to human blood group type **I** (Lacto-*N*-biose, Gal$\beta1\rightarrow3$GlcNAc) and type **II** (*N*-acetyllactosamine) carbohydrate sequences [39]. **S** determinant is found in human Tamm-Horsfall glycoprotein (THGP) and GM2. (a) **A**, **B**, **A_h**, **F_α**,**T_α** and **Tn** structural units of human ovarian cyst (HOC); (b) Asialo Sd(a^+) THGP; (c) Tri-antennary **II** and **T** units of asialo fetuin. It has three oligosaccharide side chains with two different structures *O*-glycosidically linked to Ser or Thr of the protein core as well as three carbohydrate side chains per molecule *N*-glycosidically linked to asparagine.

No. of carbohydrate chains

Gal
$\downarrow \beta 1,4$ II
GlcNAc
$\downarrow \beta 1,6$

One

Gal$\beta 1 \rightarrow$ 3GalNAc$\alpha 1 \rightarrow$ Ser/Thr

$\mathbf{T}_\alpha$

Two

Gal$\beta 1 \rightarrow$ 3GalNAc$\alpha 1 \rightarrow$ Ser/Thr,

$\mathbf{T}_\alpha$

and

$\mathbf{II}_\beta/(\mathbf{I}_\beta)$

Gal$\beta 1 \rightarrow$ 4(3) GlcNAc

$\downarrow \beta 1, 4$

Three

Gal$\beta 1 \rightarrow$ 4GlcNAc$\beta 1 \rightarrow$ 2Man$\alpha 1 \rightarrow$ 3Man$\beta 1 \rightarrow$ 4GlcNAc$\beta 1 \rightarrow$ 4GlcNAc$\beta 1$-N-Asn

$\mathbf{II}_\beta$ $\uparrow \alpha 1, 6$ C

Gal$\beta 1 \rightarrow$ 4GlcNAc$\beta 1 \rightarrow$ 2Man

$\mathbf{II}_\beta$

Branch 2   Branch 3

(2)   (3)
$\downarrow$ $\downarrow$
(7) $\rightarrow$ Gal   Gal $\leftarrow$ (9)
$\downarrow \beta 1$–4 $\mathbf{II}_\beta$ $\mathbf{II}_\beta$ $\downarrow \beta 1$–3 or 4
(8) $\rightarrow$ GlcNAc   GlcNAc $\leftarrow$ (10)
$\mathbf{I}_\beta$ $\downarrow \beta 1$–6 $\downarrow \beta 1$–6

Branch 1 (1) Gal$\beta 1 \rightarrow$ 3GlcNAc$\beta 1 \rightarrow$ 3Gal$\beta 1 \rightarrow$ 3GlcNAc$\beta 1 \rightarrow$ 3Gal$\beta 1 \rightarrow$ 3GalNAc$\alpha 1 \rightarrow$ Ser or Thr
$\uparrow$ $\uparrow$ $\uparrow \beta 1$–6
(5)   (6) $\mathbf{I}_\beta$ GlcNAc $\leftarrow$ (12)
$\mathbf{II}_\beta$ $\uparrow \beta 1$–4
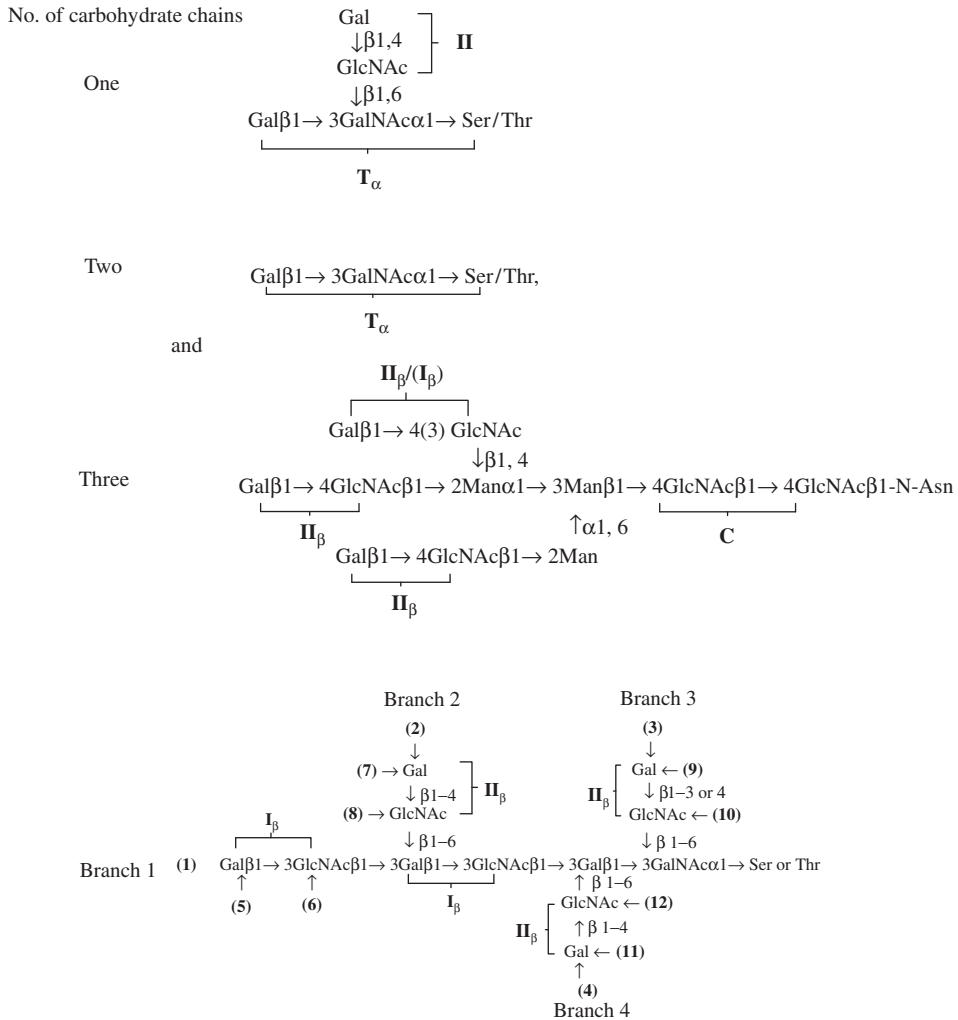Gal $\leftarrow$ (11)
$\uparrow$
(4)
Branch 4

Figure 2. (Continued) Asialo-carbohydrate chains of fetuin are shown as above [40, 41]. C: Chitin disaccharide. (d) Representative carbohydrate chains of blood group glycoprotein from human ovarian cyst fluids. The four-branched internal structure (**I–IV**) shown is the representative internal portion of the carbohydrate moiety of blood group substances to which the residues responsible for A, B, H, Le^a, and Le^b activities are attached. This structure represents precursor blood group active glycoproteins [16] and can be prepared by Smith degradation of A, B, H active glycoproteins, purified from human ovarian cyst fluids. Numbers in parentheses indicate the site of attachment for the human blood group A, B, H, Le^a, and Le^b determinants. These determinants as well as the structural units at the non-reducing end are the sources of lectin **A/A$_h$**, **B/B$_h$**, **I/II**, **T**, and **Tn** determinants in (a) [9]. A megalosaccharide of 24 sugars has not been isolated. However, most of the carbohydrate chains isolated are parts of this structure.

Blood group active glyco-protein purified from human ovarian cyst fluid	Human blood group or antigenic determinant	Sugar added	Site of addition
MSS 1st Smith[2] Beach P-1[2] Tighe P-1[2]	Ii *Pneumococcus* type 14 polysaccharide	None	
Cyst N-1 Le[a] 20% 2x	Le[a] Le[x]	LFucα1→4 LFucα1→3	(**6**) and/or (**10**) (**8**), (**10**) and/or (**12**)
Cyst JS phenol insoluble	H	LFucα1→2	(**5**), (**7**), (**9**), (**11**)
Tighe phenol insoluble	H, Le[y] H, Le[x] > Le[b]	LFucα1→2 and others as in Le[a]	(**5**), (**7**), (**9**) and/or (**11**) (**6**), (**8**), (**12**)
MSS, native MSM Cyst 9 Cyst 14	A₁ or A₂	GalNAcα1→3 and others as in H, Le[b] and Le[y]	(**1**), (**2**), (**3**) and/or (**4**) in addition to (**5**) to (**12**)
Cyst Beach phenol insoluble	B	Galα1→3 and Same as H Le[b] and Le[y]	(**1**), (**2**), (**3**) and/or (**4**), in addition to (**5**) to (**12**)

Figure 2. (Continued).

GalNAc>>>Gal (inactive) in ACL and Gal>>GalNAc (poor) in PNA. Both RCA₁₂₀ and WGAs are Galβ1→4GlcNAcβ1→6 specific, and the inhibitory profile of the monosaccharide residues with these two lectins is also very different: Gal>>>GlcNAc in RCA₁₂₀ and GlcNAcβ1→ >> Gal in wheat germ [8, 9]. From the data available, expressions of the Gal/GalNAc-specific lectins based on mammalian structural units [8–11, 22–24] are shown in Table 2.

During the past two decades, it has been established that many multi-branched oligosaccharides exhibit a significant increase in lectin-binding reactivities as compared to their linear counterparts [25, 26], "the glycoside cluster effect" in mammalian hepatic asialoglycoprotein receptors [25] is one of the important findings in this field. Effect of polyvalencies of glyco-epitopes on the binding of a lectin from the edible mushroom, *Agaricus bisporus* [27] is another important example, in which the disaccharide **II** monomer and tri-antennary Galβ1-4GlcNAc (Tri-**II**) glycopeptides, the major carbohydrate side chains of *Pneumococcus* type 14 ps, asialo fetuin and asialo human α₁-acid glycoprotein, were poor inhibitors, while their polyvalent carriers were very active, implying that there are other more complicated structural factors involved in binding besides the multi-antennary **II** sequences. To explain this phenomenon, the present concept of glycoside cluster effect has to be further defined and classified into two groups – (a) the "*Multi-antennary* or simple glycoside cluster effect" as in galactosides with hepatic lectin

Table 1. Carbohydrate structural units in mammalian glycoproteins and glycosphingolipids.

	Codes[a]	Structural Units	Sources
1	**F**	GalNAcα1→3GalNAc	Forssman pentasaccharide. Animal tissue antigens
	F$_{penta-}$	GalNAcα1→3GalNAcβ1→3Galα1→4Galβ1→4Glc	and human oncofetal glyco-epitopes, mainly in glycosphingolipids.
	F$_\alpha$	GalNAcα1→3GalNAcα1→Ser/Thr of protein core	In O-linked glycoproteins core.
	F$_\beta$	GalNAcα1→3GalNAcβ1→	Glycotope at the non-reducing end of **F**$_{penta-}$
2	**A**	GalNAcα1→3Gal	Human blood group A related **di**saccharide.
	A$_h$	GalNAcα1→3[LFucα1→2]Gal	Human blood group A related **tri**saccharide.
3	**Tn**	GalNAcα1→Ser/Thr of protein core	Tn antigen, only in O-linked glycoproteins.
4	**T**$_\alpha$	Galβ1→3GalNAcα1→Ser/Thr of protein core	The mucin-type sugar sequence on the human erythrocyte membrane.
	T$_\beta$	Galβ1→3GalNAcβ1→···ceramide	Brain glycoconjugates and gangliosides, GM$_1$.
5	**I**	Galβ1→3GlcNAc	Human blood group precursor type I and II
	I$_\beta$	Galβ1→3GlcNAcβ1	carbohydrate sequences.
6	**II**	Galβ1→4GlcNAc	Branched or linear repeated **II** sequence is part of
	II$_\beta$	Galβ1→4GlcNAcβ1→	blood group I and i epitopes. **I** and **II** are
	Tri-**II**	Tri-antennary Galβ1→4GlcNAc	precursors of ABH and Le[a], Le[b], Le[x],
	m**II**	Multivalent Galβ1→4GlcNAc	Le[y] blood group active antigens. Most of the

lectins reactive with **II** are also reactive with **I**. Lectin Tri-**II** and m**II** determinants are present at the non-reducing end of the carbohydrate chains derived from *N*- and *O*-glycans.

7	**B**	Galα1→3Gal	Human blood group B related **di**saccharide.
	B$_h$	Galα1→3[ʟFucα1→2]Gal	Human blood group B related **tri**saccharide.
8	**E**	Galα1→4Gal	Blood group P^k and P$_1$ active disaccharide. Sheep hydatid cyst glycoproteins, salivary glycoproteins of the Chinese swiftlet, glycosphingolipids in human erythrocytes, and small intestine.
9	**L**	Galβ1→4Glc	Constituent of mammalian milk.
	L$_β$	Galβ1→4Glcβ1→	Lactosyl ceramides in brain and part of carbohydrate structures in gangliosides.
10	**P**	GalNAcβ1→3Gal	Blood group P related disaccharide; glycotope at the non-reducing end of globoside.
	P$_α$	GalNAcβ1→3Galα1→	
11	**S**	GalNAcβ1→4Gal	Brain asialo-GM$_2$ disaccharide; human blood group Sd(a+) related disaccharide in most human urine secretions, Tamm-Horsfall glycoprotein.
	S$_β$	GalNAcβ1→4Galβ1→	

^{a}α, β: anomer of sugars; m: multivalent, and Tri: tri-antennary.

Table 2. Expression of binding properties of Gal/GalNAc-reactive lectins by carbohydrate structural units.

Codes	Lectins [2, 12, 13, 46, 47]	Carbohydrate Specificity
F/A	*Codium fragile* subspecies *tomentosoides* (CFA)	F_{penta-} and $T_{\alpha} > A_h \gg$ **I/II** and **L**
	Dolichos biflorus (DBA)	$F_{penta-} > A_h^a > A >$ **Tn** $\gg$ **P**
	Helix pomatia (HPA)	$F_{penta-} > A (> A_h^a) =$ **Tn, T** $\gg$ **P**
	Hog peanut (*Agaricus bisporus* agglutinin (ABA), *Amphicarpaea bracteata*)	$F_{\beta} > A \gg$ **L**
	Wistaria floribunda (WFA)	**A** $(> A_h^b)$, $F_{penta-} >$ **F/P** $>$ **Tn, I (II)** $>$ **L**
	Geodia cydonium agglutinin (GCA)	$F_{penta-}, A_h >$ **mII** $>$ **L** $>$ **II, T** $>$ **I** $\gg$ **E**
	Griffonia (Bandeiraea) simplicifolia A_4 (GSI-A_4,)	$F_{penta-} > A_h >$ GalNAc $>$ **E** $>$ **B** $>$ **I, T** $\gg$ **L, II**
F/II	*Caragana arborescens* agglutinin (CAA)	$F_{penta-} >$ **II, mTn** $>$ Sialyl **Tn**
	Wistaria sinensis agglutinin (WSA)	$F_{penta-} \geq$ **mII** $>$ **P** $>$ **II, Tn, I** and $A_h \geq$ **L/E**
A	Lima bean (LBA)	Hexa-$A_h^a > A_h^b \gg$ **B**
	Soybean (SBA, *Glycine max*)	**A** $(> A_h^b)$, **Tn** and **I (II)**
	Vicia villosa (VVA)(a mixture of A_4, A_2B_2 and B_4)	**A** $(> A_h^b)$ and **Tn** mainly
Tn	CFA	F_{penta-} and $T_{\alpha} >$ **Tn** clusters $> A_h \gg$ **I/II** and **L**
	Vicia villosa B_4 (VVL-B_4)	Two **Tn** $\gg$ one **Tn** $\gg$ one or two **T**
	Salvia sclarea (SSA)	Two **Tn** $>$ single or three sequential **Tn** structures
Tn, I/II	*Glechoma hederacea* agglutinin (GHA)	**Tn** clusters $>$ **Tn** $> A_L$, **A** $> A_h >$ **F, S, P, B** $>$ **E**
T	Peanut (PNA, *Arachis hypogaea*)	**T** $\gg$ **I (II)** $\gg$ **Tn**

T/Tn	CFA	F_α, **A**, T_α and **Tn**
	Maclura pomifera (MPA)	**T** > **Tn** >>> **I**, **II**, and **L**
	Artocarpus integrifolia (jacalin, AIA)	T_α > **P** > **T**, **Tn** >>> **I(II)** > T_β
T, Tn/II	*Bauhinia purpurea alba* (BPA)	**T** > **I(II)**, **L**, and **Tn**
	ABA	T_α and **Tn** > **I** >> GalNAc >>> **II**, **L**
	Morus nigra galactose-specific lectin (Morniga G)	T_α >> **Tn** clusters > **T** > **Tn**, **P**, Tri-**II**
T/II	*Ricinus communis* toxin (ricin, RCA$_2$)	**T** > **I(II)** and **Tn**
	Abrus precatorius agglutinin (APA)	**T** > **I/II** > **E** > **B** > **Tn**
	Sophora japonica (SJA)[c]	**T** > **I** ≥ **II** > **L**
I(II)	*Ricinus communis* agglutinin (RCA$_1$)	**II** ≥ **I** > **E**, **B** > **T**
	Datura stramonium (thorn apple, TAA)	Bi-antennary **I(II)** (penta-2,6) >> **C**[c]
	Erythrina cristagalli (coral tree, ECA)	Multi-antennary **I/II**
	Phaseolus vulgaris-L	Tri-**II**, m**II** clusters > Penta-2,6
		> Tri-2,6 > Hepta-3,6 > **II**$_\beta$ > GlcNAcβ1,2Man
	Geodia cydonium (GCL)	**A** and m**II**[d] clusters, **F** > **A** > m**II** > **L** > **II**, **T** > **I** >> **E**
B	*Griffonia (Bandeiraea) simplicifolia* B$_4$ (GSI-B$_4$)	**B** > **E** > **A**
E	*Abrus precatorius* toxin-a (Abrin-a)	**E**, **B** > **T**, **L**, **I**, and **II**
	Mistletoe lectin-I (ML-I)	**E** > **II**, **L** > **T** and **I**

[a] Substitution of Fucα1→2 to subterminal Gal is important for binding; [b] Substitution of Fucα1→2 to subterminal Gal blocks binding; [c]C, chitin disaccharide; [d]m**II**, multivalent **II**.

[25, 26] and tri-antennary **II** with a galectin from chicken liver (CG-16) [28] as well as (b) the "*High-density polyvalent* or complex glycoside cluster effect" as in macromolecular interaction of high-density **II** cluster containing glycoproteins (*Pneumococcus* type 14 ps and asialo human α_1-acid glycoproteins) with ABA. The clustered glyco-epitopes present on macromolecules generate a great enhancement in affinity with ABA up to 4.7×10^6 times as compared to their monomeric counterparts, and demonstrate the structural importance of complex carbohydrates.

To obtain a complete description of the carbohydrate specificities of the applied lectins, the following information is suggested to be necessary – (1) monosaccharide specificity (Gal, GalNAc, GlcNAc, and/or Man); (2) expression of reactivities toward mammalian disaccharide structural units or their derivatives (Table 1) in decreasing order (Table 2); (3) the most active ligand; (4) simple multivalent or cluster effect, such as glycopeptides and multi-antennary glyco-epitopes to inhibit binding; (5) complex multivalent or cluster effects present in macromolecules with known glyco-epitopes. In this report, we illustrate only, as shown in Table 2, the abilities of mammalian disaccharide structural units or their derivatives to inhibit lectin–glycan binding and express it in decreasing order. During the past years, many cases of the effect of polyvalency on binding have been found [28–35]. However, the available data are insufficient to make a solid conclusion, therefore, it is not included in this review.

Due to the multiple reactivity of lectins toward mammalian glyco-epitopes (Table 2), the possible existence of different combining sites or subsites in the same molecule has to be examined, and the differential binding properties of these combining sites (if any) have to be characterized. Recent advances in characterization, cloning, and structural analysis have allowed us to classify plant lectins into seven families of structurally and evolutionarily related proteins. Within each lectin family the overall profile and structure of the carbohydrate-binding site(s) are conserved. A closer examination of the carbohydrate specificity further indicates that most plant lectins are not targeted against plant carbohydrates but preferentially bind foreign glycans [36, 37]. Establishing the relationship between the amino acid sequences of the combining sites of plant lectins and mammalian glyco-epitopes should be an important direction to be addressed in lectinology.

3. Carbohydrate Microarrays

Our laboratories have established a high-throughput biochip platform for constructing carbohydrate microarrays [42–44]. Using this technology, carbohydrate-containing macromolecules of diverse structures, including polysaccharides, natural glycoconjugates, and mono- and oligosaccharides coupled to carrier molecules, can be stably immobilized on a glass chip without chemical modification.

This technology takes advantage of existing cDNA microarray system, including spotter and scanner, for an efficient production and application of carbohydrate microarrays.

We have demonstrated that this current platform is able to overcome a number of expected technical difficulties, by proving that (1) carbohydrate molecules can be immobilized on a nitrocellulose-coated glass slide without chemical conjugation, (2) the immobilized carbohydrates are able to preserve their immunological properties and solvent accessibility, (3) the system reaches the sensitivity, specificity, and capacity to detect a broad range of antibody specificities in clinical specimens, and (4) this technology allows highly sensitive detection, as compared to other existing technologies, of the broad range of carbohydrate–lectin/antibody interactions.

In this section, we provide a practical protocol for this high-throughput carbohydrate microarray system. We focus on an eight-chamber sub-array system that is in active use in our laboratory's routine carbohydrate research. It is our wish that the readers who have access to a standard cDNA microarray facility would be able to explore this technology for their own carbohydrate research.

3.1. Materials

3.1.1. Apparatus

    Microspotting: Cartesian Technologies' PIXSYS 5500C (Irvine, CA); or GMS 417 Arrayer, Genetic Microsystems, Inc. (Woburn, MA)
    Supporting substrate: FAST Slides (Schleicher & Schuell, Keene, NH)
    Microarray scanning: ScanArray 5000 Standard Biochip Scanning System (Packard Biochip Technologies, Inc., Billerica, MA)

3.1.2. Softwares

    Array design: CloneTracker (Biodiscovery, Inc., Marina del Rey, CA)
    Array printing: AxSys™ (Cartesian Technologies, Inc., Irvine, CA)
    Array scanning and data capturing: ScanArray Express (PerkinElmer, Torrance, CA)

3.1.3. Antibodies and lectins

    PHA-L and GSI-1 (*Phaseolus vulgaris*-L, EY Laboratories, Inc., San Mateo, CA) Streptavidin-FITC, -PE, -APC, and -Cy5 conjugates (Amersham Pharmacia, Piscataway, NJ).
    Species-specific anti-immunoglobulin antibodies and their fluorescent conjugates, APC, PE, Cy5, Alexa647, or FITC (Sigma, St. Luis, MO; BD-PharMingen, San Diego, CA; Invitogen, San Diego, CA).

3.1.4. Reagents and buffers

Dilution buffer: Saline (0.9% NaCl)
Rinsing solution: $1\times$ PBS, pH 7.4 w/0.05% (v/v) Tween 20
Blocking solution: 1% (w/v) BSA in PBS w/0.05% (w/v) NaN_3

3.2. Methods

The methods described below outline (1) design and construction of an eight-chamber sub-array system for defined purposes, (2) microspotting carbohydrate-containing molecules onto nitrocellulose-coated glass slides, (3) immunostaining and scanning of microarrays, (4) analysis of microarray data, and (5) validation of microarray findings by conventional immunological assays.

3.2.1. Design and construction of carbohydrate arrays

We have been applying an eight-chamber sub-array system to construct customized carbohydrate microarrays for defined purposes. As illustrated in Fig. 3, each microglass slide contains eight separated sub-arrays. The microarray capacity is ~500 microspots per sub-array. A single slide is, thus, designed to enable eight microarray assays. Similar sub-array designs with various array capacities are commercially available (Schleicher & Schuell, Keene, NH; ArrayIt, Sunnyvale, CA).

- Each microglass slide contains eight sub-arrays of identical content. There is chip space for 500 microspots per sub-array, with spot sizes of ~200 μm and at 300 μm intervals, center to center. A single slide is, therefore, designed to enable eight detections.
- Repeats and dilutions: We usually print carbohydrate antigens at the initial concentration of 0.1–0.5 μg/μl. The absolute amount of antigens printed on chip substrate is in the range of 0.1–0.25 ng per microspot for the highest concentration. They are further diluted at 1:3, 1:9, and 1:27. A given concentration of each preparation is repeated at least three times to allow statistic analysis of detection of identical preparation at given antigen concentration.
- Standard curves: For serological study, we include antibodies of IgG, IgA, and IgM isotype of corresponding species to serve as standard curves in microarray format. This design allows quantifying antibody signals that are captured by spotted carbohydrate antigens. In addition, such standard curves are useful for microarray data normalization and cross-chip scaling of microarray detection. For lectin study, we spot fluorescent conjugates, such as BSA conjugates of FITC, Cy5, or other dyes, to generate standard curves.
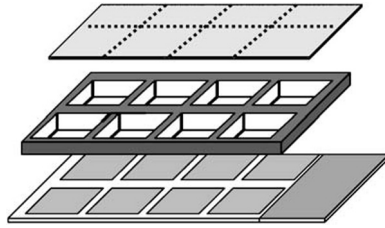
Figure 3. Illustration of the eight-chamber sub-arrays.

3.2.2. Microspotting of carbohydrates onto nitrocellulose-based substrate

Using Cartesian Technologies' PIXSYS 5500C (Irvine, CA), a high-precision robot designed for cDNA microarrays, carbohydrate antigens of various complexities are picked up by dipping quill pins into antigen/antibody solutions and printed onto nitrocellulose-coated slides in consistent amount (Schleicher & Schuell, Keene, NH; ArrayIt, Sunnyvale, CA). The complementary AxSysTM software (Cartesian Technologies, Inc., Irvine, CA) is used to instruct movement of pins about the dispense platform and the printing process.

- Prepare samples of carbohydrate antigens in 0.9% NaCl and transfer them in 96-well plates.
- Place the 96-well plates containing samples on Cartesian arrayer robot.
- Adjust program so that carbohydrate antigens are printed at spot sizes of ~150 μm and at 300 μm intervals, center to center.
- Each antigen or antibody is spotted as triplet replicates in parallel.
- The printed carbohydrate microarrays are air-dried and stored at room temperature (RT) without desiccant before application.

3.2.3. Immunostaining of carbohydrate microarrays

The staining procedure for carbohydrate microarrays is basically identical to the routine procedure for immunohistology. Immunostaining steps of carbohydrate arrays are listed below.

- Rinse printed microarray slides with $1 \times$ PBS, pH 7.4 with 0.05% Tween 20 at RT for 5 min.
- Block slides with 1% BSA in PBS containing 0.05% NaN_3 at RT for 30 min.
- Stain each sub-array with 50 μl of test sample, which is diluted in 1% BSA PBS containing 0.05% NaN_3 and 0.05% Tween 20.
- Incubate the slide in a humidified chamber at RT for 60 min.
- Wash slides by dipping them in washing buffer ($1 \times$ PBS, pH 7.4 with 0.05% Tween 20) five times with at least one-time incubation at RT for 5 min.

- Stain slides with 50 µl of titrated secondary antibodies. Anti-human (or other species) IgG, IgM, or IgA antibodies with distinct fluorescent tags, Cy5, PE or FITC, are mixed and then applied on the chips.
- Incubate the slide in a humidified chamber, protected from light at RT for 30 min.
- Wash slides five times as specified above.
- Place slide in a 50 ml falcon centrifuge tube and spin at 1,000 rpm for 5 min to remove washing buffer.
- Cover slides in a histology slide box to prevent fluorescent quenching of signal by light.

3.2.4. Microarray scanning, data processing, and statistical analysis

- Scan microarray with ScanArray Microarray Scanner (PerkinElmer Life Science) following the manufacturer's instructions.
- Fluorescence intensity values for each array spot and its background were calculated using Packard Bioscience's QuantArray software analysis packages or the updated ScanArray Express software. A staining result is considered positive if the mean fluorescent intensity value of microspot is significantly higher than the mean background of an identically stained microarray with the same fluorescent color.
- Microarray data processing and statistical analysis: a number of advanced software packages are available for microarray data normalization, statistical analysis, and pattern-recognition-based advanced data-processing (http://genome-www5.stanford.edu/resources/restech.shtml). We have been using Stanford University's Significance Analysis of Microarrays (SAM, http://www-stat.stanford.edu/~tibs/SAM/) and SAS Institute's JMP-Genomics, Proteomics, and Microarrays (http://www.jmp.com/). It is important to conceptually understand the functions of these bioinformatics tools in order to correctly interpret the results.

3.2.5. Validation of microarray observations

It is important to verify microarray findings by other experimental approaches. We usually confirm our results by at least one of the alternative immunoassays, such as ELISA, dot blot, western blot, flow cytometry, or immunohistology. Detailed examples of such investigations have been described in our recent publications [42, 44]. However, given the fact that microarray assays are highly sensitive, a positive detection in a microarray may not necessarily be reproduced by other assays that are less sensitive than microarray assays. Under such circumstances, we usually repeat the detection by microarray assays to confirm the reproducibility of a positive result. The most important validation is, however, to further characterize

the molecular targets that are discovered or suggested by microarray assays in a relevant biological system. This is discussed in the subsequent section.

4. Exploring Lectin-Binding Profiles Using Carbohydrate Microarrays

Carbohydrate microarrays share the technical advantages of miniaturized multiplex binding assays [45–51]. It is suitable for displaying a diverse panel of carbohydrates in a limited amount of chip space. It is "economical" in using carbohydrates since each saccharide is spotted on a microarray substrate in an amount (approximately sub-nanograms) that is drastically smaller than that which is required for conventional molecular or immunological assays (approximately sub-micrograms or more). Its detection sensitivity is also higher than many conventional molecular and immunological assays [43, 45, 48], which is an intrinsic property of the microarray-based miniaturized assays.

Theoretically, the improved detection sensitivity can be attributed to the fact that the binding in a microarray assay meets the so-called "ambient analyte condition" [45]. A key factor for achieving this assay condition is the use of "tiny" amounts of carbohydrate ligand in microarrays. For the microarray platform we discuss here, only sub-nanogram amounts of carbohydrates are immobilized in the microarray substrate. Binding to these miniaturized carbohydrate spots has no or minimum reduction of the concentration of lectins in the solution phase. Under such conditions, the carbohydrate–lectin interactions reach equilibrium rapidly and result in highly sensitive detection of lectins or anti-carbohydrate antibodies.

Thus, carbohydrate microarray is, in principle, well suited for measuring the relative binding reactivities of lectins with a spectrum of diverse carbohydrate structures. A few specific examples are discussed below to illustrate its application in lectin characterization.

4.1. Visualizing carbohydrate-binding profiles using carbohydrate arrays

As summarized in Section 2, the binding property of a lectin is expressed as its specific or selective reactivity with a number of glyco-epitopes that share various degrees of structural similarities or antigenic cross-reactivities. For examples, PHA-L and GS-I are Gal/GalNAc-specific lectins but have significant differences in their fine specificities (Table 2). The binding property of PHA-L is expressed as "Tri-**II** and m**II** clusters $>$ Penta-2,6 $>$ Tri-2,6 $>$ Hepta-3,6 $>$ **II**$_\beta$ $>$ GlcNAcβ1,2Man". The specificity of GS-IA$_4$ is defined as "$F_{penta-} > A_h$ (GalNAcα1$\rightarrow$2 (LFucα1$\rightarrow$3)Gal) $>$ GalNAc $>$ **E** $>$ **B** $>$ **I**, **T** $>$ **L**, **II**; GSI-B$_4$ isolectin defined as "**B** $>$ **E** $>$ **A**" (Table 2). These lectins may serve as good models for testing the potential of microarray technology in lectin characterization.

Therefore, we constructed a carbohydrate array to display a panel of 22 Gal/GalNAc-containing glycoconjugates for a "proof-of-concept" study. The eight-chamber sub-array system described in the Section 3 above was applied to produce this carbohydrate array. Each of carbohydrate preparations were spotted in four dilutions and with triplicate microspots for each dilution. This design is helpful for the quantitation and statistical analysis of microarray results. It also facilitates the graphical presentation or visualization of the binding profiles of lectins.

We assumed that if this microarray platform reached the sensitivity for measuring the lectin binding and if it reflected well the results of conventional binding assays, we would expect to visualize the lectin-binding profiles by glycan arrays. We applied lectin PHA-L and GS-I at $0.5\,\mu g/ml$ for microarray staining, which is a titrated condition for cell or tissue section staining using these reagents [44]. The images of the stained microarrays are shown in Fig. 4 (I and II) for lectin PHA-L and GS-I, respectively. The content and location of carbohydrates are listed in the table under the microarray images. A dye marker (Cy5 conjugate) was spotted in triplicates at the bottom right corner of each array.

This assay shows clearly that PHA-L is highly specific for asialoorosomucoid (ASOR) (array location C3) that expresses glyco-epitopes Tri-II and mII. Given that PHA-L has no staining of AGOR (array location A4), which is a D-galactose derivative of ASOR, the terminal non-reducing end Gal-residues are likely key elements for the specificities. In striking contrast to the highly selective binding of PHA-L, lectin GS-I shows a broad spectrum of carbohydrate-binding activities. It is reactive with seven different Gal/GalNAc-containing glycoconjugates. They are Tn-HAS, Asialo-PSM, B-dimer-BSA, Asialo-OSM, Beach ϕOH, Tij20%fr. 2nd 10%, and Beach P1 ϕOH insoluble. These carbohydrate preparations display a number of well-characterized glyco-epitopes, including Tn, T, I, i, B, and Bh.

It is interesting to compare the microarray-binding profile of GS-I with the specificities of its isolectins. GS-I is known as a tetramer consisting of subunit A and subunit B in different ratios. GS-IA$_4$ and GS-IB$_4$ have equal binding affinity for α-galactose end groups. However, the A$_4$ isolectin has a greater affinity for α-GalNAc than the B$_4$ isoform. The natural product of GS-I is a mixture of the five isolectins, A$_4$, A$_1$B$_3$, A$_2$B$_2$, A$_3$B$_1$, and B$_4$. Such molecular diversification may extend the spectrum of carbohydrate-binding profile of GS-I.

Some but not all the GS-I reactivities detected by microarray assay can be directly attributed to its subunit A and B specificities that were determined by previous studies. For example, its binding to Tn-containing carbohydrates (location C1 and C2) reflects the affinity of A subunit for α-GalNAc. Both A and B subunits are likely responsible for the binding to the terminal Gal in B, T, I, and i glyco-epitopes. However, this preparation of GS-I has no binding to cyst 9 (blood group A/location B$_4$) and Hog$_4$ (blood group A/location D$_4$). This is unexpected given that its A subunit is specific for the A-glyco-epitopes. It is also
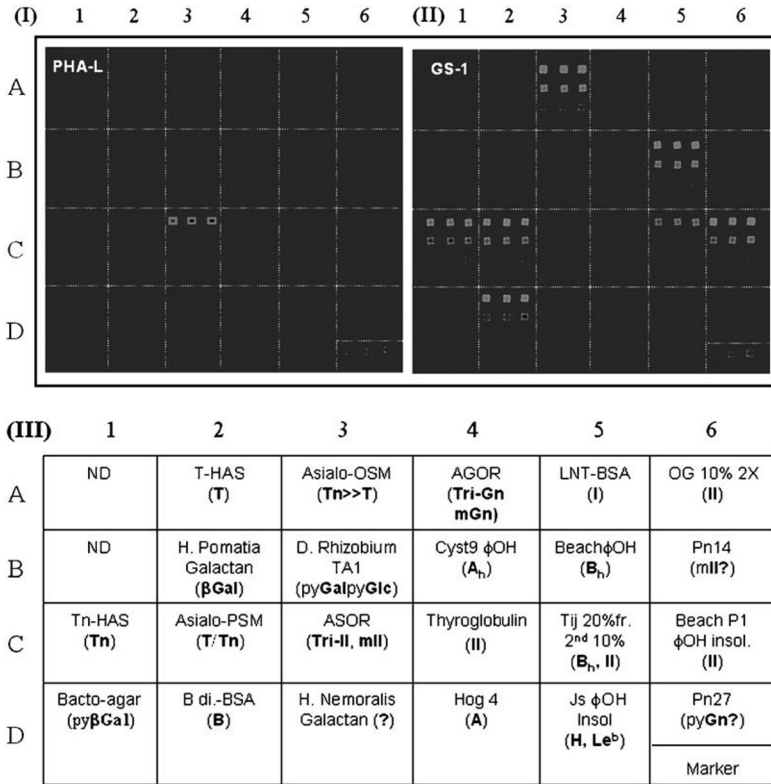
Figure 4. Carbohydrate microarrays for characterization of the carbohydrate-binding profiles of lectins. (I) An image of the microarray stained with PHA-L; (II) an image of the microarray stained with GS-I; and (III) Content and location of carbohydrate antigens in carbohydrate arrays. The carbohydrate units displayed by each carbohydrate preparation are indicated either using the codes listed in Table 1 or with its structural key elements.

not clear why this lectin is strongly reactive with asialo-OSM that displays T-glyco-epitopes but is completely negative to T-HAS. A possible interpretation of this result is that asialo-OSM displays the T-glyco-epitopes in a specific cluster configuration that substantially facilitates the binding to GS-I. By contrast, the T-glyco-epitopes of T-HAS are in the unfavorable configuration for GS-I. It is important to further characterize the specificities of individual GS-I isolectins (A_4, A_1B_3, A_2B_2, A_3B_1, and B_4). Using carbohydrate microarrays of extended structural diversity, in combination with saccharide inhibition or blocking assays, may provide more information to better understand the selective binding to specific cluster of glyco-epitopes.

4.2. Lectins as structural probes for glyco-epitope identification

Lectins and anti-carbohydrate antibodies of defined specificities are useful probes for complex carbohydrates expressed by living organisms. We present, here, an integrated strategy using these probes to facilitate the identification of immunogenic sugar moieties that are expressed by microbial pathogens. This involves three steps of experimental investigation. The first step is to perform carbohydrate microarray characterization of the antibody responses to an infectious agent or antigen preparation in order to recognize the disease- or pathogen-associated anti-carbohydrate antibody specificities. The second step focuses on identification of lectins and/or antibodies that are specific for the glyco-epitopes that are recognized by the pathogen-elicited antibodies. This provides specific structural probes to enable the third step of investigation, that is, to identify the glyco-epitopes in the candidate pathogens using specific lectins or antibodies identified by Steps 1 and 2.

We applied this strategy to explore the glyco-epitopes expressed by a previously unrecognized viral pathogen, SARS–CoV [52–54]. Our rationale was that if SARS–CoV expressed antigenic carbohydrate structures, then immunizing animals using the whole virus-based vaccines would elicit antibodies specific for these structures. In addition, if SARS–CoV displayed a carbohydrate structure that mimics host cellular glycans, then vaccinated animals may develop antibodies with autoimmune reactivity to their corresponding cellular glycans.

By characterizing the SARS–CoV neutralizing antibodies elicited by an inactivated SARS–CoV vaccine, we detected autoantibody reactivity specific for the carbohydrate moieties of an abundant human serum glycoprotein ASOR [44]. This surveillance provides important clues for the selection of specific immunologic probes to further examine whether SARS–CoV expresses antigenic structures that imitate the host glycan. As shown in the Fig. 4 above, lectin PHA-L is specific for glyco-epitopes Tri-**II** or m**II** of ASOR. Using PHA-L as a structural probe, we confirmed that SARS–CoV expresses the PHA-L reactive antigenic structure (see Ref. [44] for details). We, thus obtained, immunological evidence that a carbohydrate structure of SARS–CoV shares antigenic similarity with host glycan complex carbohydrates. This viral component is probably responsible for the stimulation of the autoantibodies directed at a cellular glycan complex carbohydrate.

These observations raise important questions about whether autoimmune responses are in fact elicited by SARS–CoV infection and whether such autoimmunity contributes to SARS pathogenesis. ASOR is an abundant human serum glycoprotein and the ASOR-type complex carbohydrates are also expressed by other host glycoproteins [55, 56]. Thus, the human immune system is generally non-responsive to these "self" carbohydrate structures. However, when similar sugar moieties were expressed by a viral glycoprotein, their cluster configuration could differ significantly from those displayed by a cellular glycan, and in this manner generate a novel "non-self" antigenic structure.

A documented example of such antigenic structure is a broad-range HIV-1 neutralization epitope recognized by a monoclonal antibody 2G12. This antibody is specific for a unique cluster of sugar chains displayed by the gp120 glycoprotein of HIV-1 [57]. It is, therefore, important to examine whether naturally occurring SARS–CoV expresses the Tri-**II** or m**II**-type autoimmune reactive sugar moieties. During SARS epidemic, the viruses replicate in human cells. Their sugar chain expression may differ from the monkey cell produced viral particles. Scanning of the serum antibodies of SARS patients using glycan arrays or other specific immunologic tools may provide information that can shed light on this question.

In summary, the recent establishment of carbohydrate-based microarrays, and especially the availability of different technological platforms to meet the multiple needs of carbohydrate research [42, 58–64], marks an important developmental stage of postgenomic research. It is our prediction that microarray-based broad-range binding assays may substantially extend the scope of current carbohydrate research, including lectin characterization and classification. We have described here a practical platform of carbohydrate microarrays and discussed a few examples to illustrate the application of this technology in lectin characterization. We also present an integrated experimental strategy to facilitate the identification of the glyco-epitopes in biological systems, such as a newly emerged viral pathogen, SARS–CoV. This research strategy is likely applicable for the exploration of complex carbohydrates that are differentially expressed by host cells, including human cancers and stem cells at various stages of differentiation.

References

[1] S.A. Brooks, M.V. Dwek and U. Schumacher, Functional & Molecular Glycobiology, BIOS Scientific Publishers Ltd, Oxford, 2002.

[2] H. Feinberg, D.A. Mitchell, K. Drickamer and W.I. Weis, Science, 294 (2001) 2163–2166.

[3] T. Feizi, Adv. Exp. Med. Biol., 152 (1982) 167–177.

[4] T. Feizi, Trends. Biochem. Sci., 19 (1994) 233–234.

[5] P.R. Crocker and T. Feizi, Curr. Opin. Struct. Biol., 6 (1996) 679–691.

[6] S.I. Patenaude, S.M. Vijay, Q.L. Yang, H.J. Jennings and S.V. Evans, Acta Crystallogr. D. Biol. Crystallogr., 54 (1998) 1005–1007.

[7] S.I. Patenaude, C.R. MacKenzie, D. Bilous, R.J. To, S.E. Ryan, N.M. Young and S.V. Evans, Acta Crystallogr D Biol. Crystallogr., 54 (1998) 1456–1459.

[8] A.M. Wu, S.C. Song, M.S. Tsai and A. Herp, Adv. Exp. Med. Biol., 491 (2001) 551–585.

[9] A.M. Wu and S.J. Sugii, Adv. Exp. Med. Biol., 228 (1988) 205–263.

[10] A.M. Wu, S.C. Song, S. Sugii and A. Herp, Indian J. Biochem. Biophys., 34 (1997) 61–71.

[11] A.M. Wu, Mol. Cell. Biochem., 61 (1984) 131–141.

[12] G.F. Springer, Prog. Allergy, 15 (1971) 9–77.

[13] A. Makita, C. Suzuki, Z. Yosizawa and T. Konno, Tohoku J. Exp. Med., 88 (1966) 277–288.

[14] S. Hakomori and R. Kannagi, J. Natl. Cancer Inst., 71 (1983) 231–251.

[15] A.M. Wu, Adv. Exp. Med. Biol., 228 (1988) 351–394.

[16] W.M. Watkins. In: A. Gottschalk, (Ed.), Glycoproteins, Elsevier, Amsterdam, 1972, pp. 830–890.

[17] W.M. Watkins, Adv. Hum. Genet., 10 (1980) 1–136, 379–385.

[18] F. Maisonrouge-McAuliffe and E.A. Kabat, Arch. Biochem. Biophys., 175 (1976) 71–80.

[19] M. Fukuda and M.N. Fukuda, Biology of glycoproteins, Plenum Press, New York, 1984.

[20] J.M. Wieruszeski, J.C. Michalski, J. Montreuil, G. Strecker, J. Peter-Katalinic, H. Egge, H. Van Halbeek, J.H. Mutsaers and J.F. Vliegenthart, J. Biol. Chem., 262 (1987) 6650–6657.

[21] A.M. Wu, Adv. Exp. Med. Biol., 491 (2001) 55–64.

[22] J.T. Gallagher, Biosci. Rep., 4 (1984) 621–632.

[23] O. Mäkelä, Ann. Med. Exp. Biol. Fenn. Suppl., 1135 (1957) 1–156.

[24] I. Goldstein and R. Poretz, In: I. Liener, N. Sharon and I. Goldstein, (Eds.), The lectins, properties, functions and applications in biology and medicine, Academic Press, New York, 1986, pp. 33–247.

[25] R.T. Lee and Y.C. Lee, Glycoconj. J., 17 (2000) 543–551.

[26] Y.C. Lee, FASEB J., 6 (1992) 3193–3200.

[27] A.M. Wu, J.H. Wu, A. Herp and J.H. Liu, Biochem. J., 371 (2003) 311–320.

[28] A.M. Wu, J.H. Wu, M.S. Tsai, H. Kaltner and H.J. Gabius, Biochem. J., 358 (2001) 529–538.

[29] A. Wu, J. Wu, L. Lin, S. Lin and J. Liu, Life Sci., 72 (2003) 2285–2302.

[30] A.M. Wu, J.H. Wu, J.H. Liu and T. Singh, Life Sci., 74 (2004) 1763–1779.

[31] A.M. Wu, FEBS Lett., 562 (2004) 51–58.

[32] A.M. Wu, J.H. Wu, J.H. Liu, T. Singh, S. Andre, H. Kaltner and H.J. Gabius, Biochimie, 86 (2004) 317–326.

[33] T. Singh, U. Chatterjee, J.H. Wu, B.P. Chatterjee and A.M. Wu, Glycobiology, 15 (2005) 67–78.

[34] T. Singh, J.H. Wu, W.J. Peumans, P. Rouge, E.J. Van Damme, R.A. Alvarez, O. Blixt and A.M. Wu, Biochem. J., 393 (2006) 331–341.

[35] A.M. Wu, T. Singh, J.H. Wu, M. Lensch, S. Andre and H.J. Gabius, Glycobiology, 16 (2006) 524–537.

[36] W.J. Peumans, E.J. Van Damme, A. Barre and P. Rouge, Adv. Exp. Med. Biol., 491 (2001) 27–54.

[37] W. Peumans, A. Barre, G. Hao, P. Rouge and E. Van Damme, Trends Glycosci. Glycotechnol., 12 (2000) 83–101.

[38] A. Makita, C. Suzuki and Z. Yosizawa, J. Biochem., 60 (1966) 502–513.

[39] A.M. Wu, A. Herp, S.C. Song, J.H. Wu and K.S. Chang, Life Sci., 57 (1995) 1841–1852.

[40] B. Nilsson, N.E. Norden and S. Svensson, J. Biol. Chem., 254 (1979) 4545–4553.

[41] R.G. Spiro and V.D. Bhoyroo, J. Biol. Chem., 249 (1974) 5704–5717.

[42] D. Wang, S. Liu, B.J. Trummer, C. Deng and A. Wang, Nat. Biotechnol., 20 (2002) 275–281.

[43] D. Wang, Proteomics, 3 (2003) 2167–2175.

[44] D. Wang and J. Lu, Physiol. Genomics, 18 (2004) 245–248.

[45] R. Ekins, F. Chu and E. Biggart, Ann. Biol. Clin., 48 (1990) 655–666.

[46] A. Lueking, M. Horn, H. Eickhoff, K. Bussow, H. Lehrach and G. Walter, Anal. Biochem., 270 (1999) 103–111.

[47] G. MacBeath and S.L. Schreiber, Science, 289 (2000) 1760–1763.

[48] D. Stoll, M.F. Templin, M. Schrenk, P.C. Traub, C.F. Vohringer and T.O. Joos, Front. Biosci., 7 (2002) C13–C32.

[49] P.J. Utz, Immunol. Rev., 204 (2005) 264–282.

[50] W.H. Robinson, C. DiGennaro, W. Hueber, B.B. Haab, M. Kamachi, E.J. Dean, S. Fournel, D. Fong, M.C. Genovese, H.E. de Vegvar, K. Skriner, D.L. Hirschberg, R.I. Morris, S. Muller, G.J. Pruijn, W.J. Van Venrooij, J.S. Smolen, P.O. Brown, L. Steinman and P.J. Utz, Nat. Med., 8 (2002) 295–301.

[51] W.H. Robinson, P. Fontoura, B.J. Lee, H.E. de Vegvar, J. Tom, R. Pedotti, C.D. DiGennaro, D.J. Mitchell, D. Fong, P.P. Ho, P.J. Ruiz, E. Maverakis, D.B. Stevens, C.C. Bernard, R. Martin, V.K. Kuchroo, J.M. Van Noort, C.P. Genain, S. Amor, T. Olsson, P.J. Utz, H. Garren and L. Steinman, Nat. Biotechnol., 21 (2003) 1033–1039.

[52] R.A. Fouchier, T. Kuiken, M. Schutten, G. Van Amerongen, G.J. Van Doornum, B.G. Van Den Hoogen, M. Peiris, W. Lim, K. Stohr and A.D. Osterhaus, Nature, 423 (2003) 240.

[53] T.G. Ksiazek, D. Erdman, C.S. Goldsmith, S.R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J.A. Comer, W. Lim, P.E. Rollin, S.F. Dowell, A.E. Ling, C.D. Humphrey, W.J. Shieh, J. Guarner, C.D. Paddock, P. Rota, B. Fields, J. DeRisi, J.Y. Yang, N. Cox, J.M. Hughes, J.W. LeDuc, W.J. Bellini and L.J. Anderson, N. Engl. J. Med., 348 (2003) 1953–1966.

[54] P.A. Rota, M.S. Oberste, S.S. Monroe, W.A. Nix, R. Campagnoli, J.P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M.H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J.L. DeRisi, Q. Chen, D. Wang, D.D. Erdman, T.C. Peret, C. Burns, T.G. Ksiazek, P.E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rassmussen, R. Fouchier, S. Gunther, A.D. Osterhaus, C. Drosten, M.A. Pallansch, L.J. Anderson and W.J. Bellini, Science, 300(5624) (2003) 1394–1399.

[55] R.D. Cummings and S. Kornfeld, J. Biol. Chem., 259 (1984) 6253–6260.

[56] F. Pacifico, N. Montuori, S. Mellone, D. Liguoro, L. Ulianich, A. Caleo, G. Troncone, L.D. Kohn, B. Di Jeso and E. Consiglio, Mol. Cell. Endocrinol., 208 (2003) 51–59.

[57] D.A. Calarese, C.N. Scanlan, M.B. Zwick, S. Deechongkit, Y. Mimura, R. Kunert, P. Zhu, M.R. Wormald, R.L. Stanfield, K.H. Roux, J.W. Kelly, P.M. Rudd, R.A. Dwek, H. Katinger, D.R. Burton and I.A. Wilson, Science, 300 (2003) 2065–2071.

[58] W.G. Willats, S.E. Rasmussen, T. Kristensen, J.D. Mikkelsen and J.P. Knox, Proteomics, 2 (2002) 1666–1671.

[59] F. Fazio, M.C. Bryan, O. Blixt, J.C. Paulson and C.H. Wong, J. Am. Chem. Soc., 124 (2002) 14397–14402.

[60] S. Fukui, T. Feizi, C. Galustian, A.M. Lawson and W. Chai, Nat. Biotechnol., 20 (2002) 1011–1017.

[61] B.T. Houseman and M. Mrksich, Chem. Biol., 9 (2002) 443–454.

[62] S. Park and I. Shin, Angew. Chem. Int. Ed. Engl., 41 (2002) 3180–3182.

[63] E.W. Adams, D.M. Ratner, H.R. Bokesch, J.B. McMahon, B.R. O'Keefe and P.H. Seeberger, Chem. Biol., 11 (2004) 875–881.

[64] G.T. Carroll, D. Wang, N.J. Turro and J.T. Koberstein, Langmuir, 22 (2006) 2899–2905.