NCBI Bookshelf. A service of the National Library of Medicine, National Institutes of Health.

Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009.

Chapter 27 Principles of Glycan Recognition

Richard D Cummings and Jeffrey D Esko.

Glycans interact with many types of proteins, as exemplified by their binding to enzymes, antibodies, and glycan-binding proteins (GBPs). Binding of glycans to proteins represents the major way in which the information contained in glycan structures is recognized, deciphered, and put into biological action. This chapter describes the structural and thermodynamic principles of glycan–GBP interactions and the methods used to measure association constants.

GLYCAN–PROTEIN RECOGNITION

The general classes of GBPs are described in Chapter 26. GBPs differ in the types of glycans they recognize and in their binding affinity. A common question is: How does a GBP (such as a bacterial adhesin, toxin, plant lectin, viral hemagglutinin, antibody, or animal GBP) bind to only a very limited number of glycans (or even a single glycan) among the thousands that are produced by a cell? A fundamental question at the molecular level is: What forces allow a specific glycan to leave the aqueous phase to enter a protein-combining site? Understanding the molecular basis of the interaction between glycans and proteins is not merely an academic exercise in chemistry and physics, it is essential for the development of a real molecular appreciation of how glycoconjugates affect fundamental biological responses. Furthermore, the information that emerges from these kinds of studies provides a springboard for designing specifically tailored compounds that inhibit a glycan-binding site. For example, understanding the principles of glycan–protein interactions led to the development of high-affinity inhibitors of viral neuraminidase, which is involved in the removal of sialic acids from host glycoprotein receptors, and these inhibitors have proven utility in reducing the duration and spread of influenza infection (see Chapters 39, 50, and 51).

HISTORICAL BACKGROUND

Many of the early studies of glycan–protein interactions focused on the recognition of glycans by enzymes, including the endoglycosidase lysozyme (which can degrade bacterial cell walls) and certain enzymes involved in intermediary metabolism (such as glycogen and starch synthases and phosphorylases). Following their discovery of the antibacterial action of lysozyme and penicillin, Sir Alexander Fleming, Ernst Chain, and Sir Howard Florey received the Nobel Prize in 1945. Lysozyme was subsequently shown to be a highly specific endoglycosidase capable of cleaving α1-4-linkages in bacterial peptidoglycan (see Chapter 20). Glycogen synthase was found to generate α1-4-glucosyl residues in glycogen, whereas other branching and debranching enzymes recognized α1-6-branched glucose residues. Studies on the metabolism of glycogen by these enzymes led to the award of the Nobel Prize in 1970 to Luis F. Leloir for his discovery of sugar nucleotides and their role in the biosynthesis of glycans.

The concept of glycans being specifically recognized by proteins dates back to Emil Fischer, who used the phrase “lock and key” to refer to enzymes that recognize specific glycan substrates. Lysozyme was the first “carbohydrate-binding protein” to be crystallized, and determination of its three-dimensional structure led to the realization that specific interactions occur between sugars and proteins. The structure of lysozyme was solved in a complex with a tetrasaccharide in an elegant series of studies in the late 1960s. Lysozyme is an ellipsoid protein, which has a long

cleft that runs for most of the length of one surface of the molecule. This cleft is astonishingly large, considering that lysozyme has only 129 amino acids, and it is capable of accommodating a hexasaccharide and cleaving it into di-and tetrasaccharide products. Other GBPs whose three-dimensional structures are of historical significance are concanavalin A (ConA; crystal structure reported in 1972) and influenza virus hemagglutinin (crystal structure reported in 1981). Much of the initial work important to our understanding of glycan–protein interactions was gathered in studies on the combining sites of plant lectins and antibodies toward specific blood group antigens. From these and many recent studies, it is now clear that the specific recognition of sugars by proteins occurs by multiple mechanisms.

BIOPHYSICAL METHODS TO DETERMINE GLYCAN–PROTEIN INTERACTIONS

Two of the most widely used biophysical approaches for examining protein–glycan interactions at the molecular level are X-ray crystallography and nuclear magnetic resonance (NMR). Several hundred three-dimensional structures of GBPs are listed in various structural protein data banks (e.g., Research Collaboratory for Structural Bioinformatics [RCSB], Protein Data Bank [PDB], and the National Center for Biotechnology Information [NCBI]). Additionally, several thousand three-dimensional structures of mono- and oligosaccharides are stored in the Cambridge Structural Database.

A resolution of at least 2–2.5 Å is required to identify accurately the positions and mode of binding of glycans. However, such resolution is often difficult to obtain. In some cases, a crystal structure can be obtained for a GBP independent of ligand, and its structure and potential binding site can be predicted using information from the three-dimensional structure of a homologous protein. Since small molecules often cocrystallize with a GBP better than large molecules, a lot of our knowledge about glycan–GBP interactions at the atomic level is based on co-crystals of GBPs with unnatural ligands. Thus, a great challenge exists in attempting to understand glycan–GBP interactions in the context of natural glycans present as a glycoprotein, glycolipid, or proteoglycan.

In NMR, the proton–proton distances in small molecules (generally <2000 daltons) can be obtained following assignment of the proton resonances through multidimensional techniques, such as measuring the nuclear Overhauser effect (NOE). This information coupled with computational methods that employ computer modeling allows the prediction of the free-state glycan conformation in solution. NMR studies of free glycans have shown that monosaccharides are relatively rigid molecules in solution, but an oligosaccharide actually has a higher degree of flexibility because of the general freedom of rotation about glycosidic bonds (see Chapter 2). The main approaches for using NMR to define the conformation of bound glycans are NOE spectroscopy (NOESY), transfer NOE (trNOESY; transferred rotating-frame Overhauser effect spectroscopy), and heteronuclear single-quantum correlation (HSQC). Computer-assisted modeling strategies and information from glycan solution conformations and protein three-dimensional structures can be combined with NMR to provide even more information about the molecular details of the interactions between glycans and proteins. Although these approaches are highly informative, they are limited by the degree to which small glycans structurally mimic the larger macromolecule to which they are usually attached.

ATOMIC DETAILS OF PROTEIN–GLYCAN INTERACTIONS

Several hundred structures of glycan–protein complexes have been solved by X-ray crystallography and NMR spectroscopy. In most cases, the glycan-binding sites typically accommodate one to four sugar residues, although in some complexes, recognition extends over larger numbers of residues and may include aglycone components, such as

the peptide or lipids to which the glycans are attached. The complex of cholera toxin, a pentameric protein, bound to the ganglioside GM1 pentasaccharide illustrates this general pattern (Figure 27.1). Cholera toxin is a soluble protein generated by *Vibrio cholerae* that has an AB_5 sub-unit structure, where the single A subunit (~22 kD) is an enzyme (ADP-ribosyltransferase) and the five identical B subunits (~55 kD) each bind GM1 (see Chapter 34). Only three of the five monosaccharides of each glycolipid ligand actually make contact with each B subunit. The terminal β-linked galactose residue is almost completely buried in the protein-combining site, making numerous hydrogen-bonding interactions with protein functional groups, some via bound water molecules. In addition, an aromatic amino acid (Trp-88) stacks against the hydrophobic "underside" of one of the sugar residues, in this case, the terminal β-linked galactose residue. Hydrophobic interactions are very common in glycan–protein complexes and can involve aromatic residues as well as alkyl side chains of amino acids in the combining site. The penultimate β-linked GalNAc residue makes only minor contacts with the protein, whereas the Galβ1-4Glc reducing-end disaccharide makes no contact at all. About 350 Å of the solvent-accessible surface of the pentasaccharide is buried in the B subunit. About 43% of this area is due to interactions with the sialic acid, 39% is due to galactose, and 17% is due to *N*-acetylgalactosamine.

Knowing the three-dimensional structure of a glycan-protein complex can reveal much about the specificity of binding, changes in conformation that take place on binding, and the contribution of specific amino acids to the interaction. However, one would also like to determine the affinity of the interaction. Since the forces involved in the binding of a glycan to a protein are the same as for the binding of any ligand to its receptor (hydrogen bonding, electrostatic or charge interactions, van der Waals interactions, and dipole attraction), it is tempting to try to calculate their contribution to overall binding energy, which can then be related to an affinity constant (K_a). Unfortunately, calculating the free energy of association is difficult for several reasons, including problems in defining the conformation of the unbound versus the bound glycan, changes in bound water within the glycan and the binding site (as discussed below), and conformational changes in the GBP upon binding.

The "Water Problem"

As shown in Figure 27.2, the overall process of binding typically involves the union of a hydrated polyhydroxylated glycan and a hydrated protein-combining site. If a surface on the glycan is complementary to the protein-combining site, water can be displaced and binding occurs. When the complex finally forms, it presents a new surface to the surrounding medium, which will also be hydrated. Solvation/desolvation energies are very large because of entropy from the disordering of water molecules, but their contribution to binding cannot be reliably determined with existing models. Furthermore, glycans may undergo a conformational change upon binding, changing their internal energy and solvation. Thus, although one can estimate the energetic contributions of van der Waals and hydrogen bonding interactions in the combining site, errors in the estimation of the attendant solvation energy changes render the overall calculations of binding energy problematic.

Valency

Another complicating issue concerns valency, which greatly increases the association of GBPs with their glycan ligands under biological conditions. The oligosaccharide of ganglioside GM1 is unusual in that it binds strongly to the B subunits of cholera toxin with high affinity because of the specific and multiple interactions in the combining site (K_d of ~40 nM). However, the affinity of most single glycan–protein interactions is generally low (mM–μM K_d values). In nature, many GBPs are oligomeric or may be membrane-associated proteins, which allows aggregation of the GBP in the plane of the membrane. Many of the glycan ligands for GBPs are also multivalent. In the case of cholera toxin, five B subunits present in the holotoxin interact with five molecules of GM1 normally present in the

cell membrane. The interaction of multiple subunits with a multivalent display of GM1 raises the affinity of the interaction by several orders of magnitude (K_d of ~40 pM). The term *avidity* is used to refer to the strength of multivalent ligand binding.

The density of binding sites on the ligand can also affect the affinity of binding. Ligands for some GBPs may be glycoproteins that carry one or more multiantennary N-linked chains. Mucins present potentially hundreds of glycans, and their proximity can affect conformation and presentation of the ligands. Some polysaccharides, such as glycosaminoglycans, have multiple binding sites located along a single chain (see Chapter 16). In addition, nonglycan components, such as tyrosine sulfate, lipid, or peptide determinants, may also cooperate with glycans to provide relatively high affinity and specific interactions. A good example of this is the interaction of PSGL-1 (P-selectin glycoprotein ligand-1) with P- and L-selectin (see Chapter 31), where the combination of sialyl Lewis^x on a core-2 O-glycan coupled with multiple adjacent tyrosine sulfate and amino acid residues increases the binding affinity manyfold, from low affinity (K_d ~ 0.1 mM) to a physiologically relevant level (K_d ~ 0.1–0.3 μM).

THERMODYNAMICS OF BINDING

The interaction of glycans with GBPs can also be described thermodynamically and kinetically. The binding of a lectin (L) to a glycan (G) is governed by Equation 1. The affinity constant, K, is defined as an association constant (or K_a) by Equation 2 and is equal to k_1/k_2. Like any equilibrium constant, K is related to the standard free-energy change of the binding reaction at pH 7 (ΔG_o) in kcal per mole, as shown by Equation 4, where R is the gas constant (0.00198 kcal/mol-degree) and T is the absolute temperature (298°K). The affinity constant K is related to the thermodynamic parameters ΔG, ΔH, and ΔS (see Equation 4), which represent the changes in free energy, enthalpy, and entropy of binding, respectively.

$$\text{Lectin (L)} + \text{Glycan (G)} \underset{K_2}{\overset{K_1}{\rightleftarrows}} \text{LG} \quad (1)$$

$$K_a = [\text{LG}]/[\text{L}][\text{G}] = k_1/k_2 \quad (2)$$

$$K_d = [\text{L}][\text{G}]/[\text{LG}] = 1/K_a = k_2 k_1 \quad (3)$$

$$\Delta G_o = RT \ln K = \Delta H - \text{T}\Delta S \quad (4)$$

The rate constant k_1 is expressed in units of $M^{-1}sec^{-1}$ or $M^{-1}min^{-1}$, whereas k_2 is expressed in units of sec^{-1} or min^{-1}. Although it is important to define K_a, k_1, k_2, ΔG, ΔH, and ΔS for each binding phenomenon under consideration, investigators often discuss data in terms of the $K_d = 1/K_a$ (see Equation 3), because the units are in concentration (millimolar, micromolar, nanomolar, etc.).

Binding of a monovalent GBP to a monovalent ligand is easily defined by the equilibrium kinetics described in the

above equations. However, with multivalent ligands or GBPs, multiple affinities occur and a more complex binding equilibrium (more accurately described by a set of equilibrium constants) must be used. Typically, for multivalent ligands and GBPs, the values reported for affinity are apparent affinity constants and usually measure the avidity.

There are many different ways to study binding of glycans to proteins, and each approach has its advantages and disadvantages in terms of thermodynamic rigor, amounts of protein and glycan needed, and the speed of analysis. Below is a discussion of some of the major ways in which the binding between a glycan and protein can be studied.

TECHNIQUES TO STUDY GLYCAN–PROTEIN INTERACTIONS

Much of the available information about glycan–protein interactions derives from studies of relatively small glycan ligands interacting with a protein. In examining these interactions, two broad categories of techniques have been applied: (1) kinetic and near-equilibrium methods, such as equilibrium dialysis and titration calorimetry; and (2) nonequilibrium methods such as glycan microarray screening, hapten inhibition, ELISA-based approaches, and agglutination.

Kinetics and Near-Equilibrium Methods

Equilibrium Dialysis for Measuring K_d Values and Interaction Valency

The simple concept of equilibrium dialysis involves placing a solution of a GBP (e.g., a lectin or an antibody) in a dialysis chamber that is permeable to a glycan or other small hapten. In **glycobiology**, the term hapten, which was coined by Karl Landsteiner in his study of immunity to antigens, is often used to denote a small glycan that competitively binds to a lectin and competes for its binding to a more complex ligand. Thus, it can be used as a surrogate ligand. The chamber is then placed in a known volume of buffer that contains the glycan in the concentration range of the expected K_d. At equilibrium, the concentration of bound plus free glycan inside the bag [In] and free glycan outside the bag [Out] will depend on the concentration and affinity of the GBP inside the bag. From this information, both the K_a and the valence n can be determined from the relationship shown in Equation 5, where r is the molar ratio of glycan bound to GBP; c is the concentration of unbound glycan [Out]. The amount of glycan bound is determined by simply subtracting [Out] from [In].

$$r/c = K_a n - K_a r \qquad 5$$

A plot of r/c versus r for different hapten concentrations is known as a Scatchard plot and will approximate to a straight line with a slope of $-K_a$. The valence of binding (number of binding sites per mol) is defined by the r intercept at an infinite hapten concentration. If such an analysis were done with cholera toxin, for example, one would obtain five binding sites per mol of AB_5 complex, or one mol per mol of B subunit.

As in any technique for determining binding constants, a number of important assumptions are made and their validity must be considered. These include demonstrating that the protein and its hapten are stable and active over the course of the experiment, the hapten is freely diffusible, the complex is at equilibrium, and structurally unrelated haptens, not expected to bind, show no apparent binding in the experimental setup. The following are several advantages to equilibrium dialysis: (1) The ease of the approach and sophisticated equipment is not needed; (2) if the affinity is high, then relatively small amounts of protein are needed (≤1 mmole); (3) if the affinity is high, only small amounts of

hapten may be required; (4) if the protein and haptens are very stable, they may be recovered and reused; (5) radioactive haptens may be used; and (6) reliable equilibrium measurements can be made. Some drawbacks of the approach are that (1) it can only provide K_a; (2) if the affinity of the GBP or antibody for the hapten is low, then relatively large amounts of both may be required; and (3) many different measurements must be made and this may require many days or weeks to complete.

A variation of this technique is illustrated by the equilibrium gel-filtration method developed by Hummel and Dreyer. In the Hummel–Dreyer method, a GBP is applied to a gel-filtration column that has been preequilibrated with a glycan of interest that is easily detectable (e.g., by radioactive or fluorescent tagging). As the protein binds to the ligand, a complex is formed that emerges from the column as a "peak" above the baseline of ligand alone, followed by a "trough" (where the concentration of ligand is decreased below the baseline) that extends to the included or salt volume of the column. The amount of complex formed is easily determined by the known specific activity of the ligand. Because the amount of complex formed is directly proportional to the amount of protein (or ligand) applied, it is easy to calculate a binding curve from several different Hummel–Dreyer column profiles at different concentrations of either protein or ligand. This binding curve allows the calculation of the equilibrium constant of the interaction. The advantages and drawbacks of this technique are generally the same as for equilibrium dialysis, except that Hummel–Dreyer analyses are often quicker to perform and can be used with ligands of many different sizes. Such an approach has been invaluable in defining the equilibrium binding of selectins to their ligands (see Chapter 31).

Affinity Chromatography to Assess the Specificity of Binding

In many cases, affinity chromatography is simply used to identify interacting partners. In this technique, a GBP is immobilized to an affinity support, such as Affi-Gel™, CNBr-activated Sepharose, Ultralink™, or some other activated support. If a glycan binds tightly to an immobilized GBP, buffer containing a known hapten may be added to force dissociation of the complex. For example, oligomannose-type and hybrid-type N-glycans will bind avidly to an agarose column containing the plant lectin concanavalin A (ConA-agarose) and 10–100 mM α-methyl manno-side is required to elute the bound material efficiently. In contrast, many highly branched complex-type N-glycans will not bind. Biantennary complex-type N-glycans bind to ConA-agarose, but they do not bind as tightly as high-mannose-type N-glycans and their elution can be effected by using 10 mM α-methyl glucoside. In this manner, one can assess the binding specificity of a GBP. In practice, this approach is rather crude, and although it gives valuable practical information about the capacity of an immobilized lectin to bind specific glycans, it does not provide quantitative affinity measurements. A variant of this method is to immobilize the glycan ligand through covalent linkage or by capturing a biotinylated glycan on a streptavidin-linked surface and then measuring GBP binding.

A more sophisticated version of this approach, termed frontal affinity chromatography, does provide quantitative measurements of the equilibrium binding constants. In this technique, a solution containing a glycan of known concentration is continuously applied to a column of immobilized GBP, and the elution front of the glycan from the column is monitored. Eventually, enough ligand is added through continuous addition that its concentration in the eluant equals that in the starting material. If the glycan has no affinity for the GBP, it will elute in the void volume V_o; if, however, the glycan interacts with the GBP, it will elute after the V_o and at a volume V_f (Figure 27.3).

The advantages of frontal affinity chromatography are similar to those discussed for equilibrium dialysis: (1) The approach is easy and inexpensive; (2) if the affinity is high, then relatively small amounts of protein are needed (≤1 mmole), and only a single column is required; (3) correspondingly, small amounts of glycan may be used if the K_d is in the range of 10 nM to 10 mM; (4) if the glycans are stable, they may be recovered and reused; (5) radioactive

glycans may be used; and (6) reliable equilibrium measurements can be made. There are several drawbacks to this approach, including (1) only the K_d can be derived, not k_{on} or k_{off}; (2) the conjugation of the GBP to the matrix must be stable and the protein must retain reasonable activity for many different column runs; (3) many different column runs must be made with a single glycan; and (4) if the K_d is high (>1 mM), this approach is not really feasible. Overall, frontal affinity chromatography is quite useful and has now been automated.

Titration Calorimetry to Measure K_d and Binding Enthalpy

This is one of the most rigorous means of defining the equilibrium binding constant between a glycan and a GBP. The binding of a glycan to the GBP is measured as a change in enthalpy through isothermal titration microcalorimetry using a commercial microcalorimeter. In this technique, a solution containing a glycan of interest is added in increments into a solution containing a fixed concentration of GBP. The glycan is added at many intervals and the heat evolved from binding is measured relative to a reference cell. Over the course of the experiment, the concentration of glycan is increased in the mixing cell over a glycan-to-GBP molar ratio of 0–10. The change of heat capacity of binding is determined and the data are replotted as kcal/mole of injectant versus the molar ratio (Figure 27.4). These data are then analyzed by replotting data to obtain the K_d. The heat change is directly related to the enthalpy of reaction ΔH. From knowledge of the K_d and ΔH, and using Equation 4, it is possible to define the binding entropy ΔS.

The major advantage of this approach is that it can provide all major thermodynamic information about the binding of a glycan to a GBP, and thus, it is highly superior to equilibrium dialysis and affinity chromatography. The disadvantages of this approach are that (1) relatively large amounts of protein may be required to conduct multiple experiments (>10 mg); (2) relatively large amounts of glycans may be required; and (3) because of the above-mentioned problem, it is not typical for such analyses to use a wide range of different glycans. Nevertheless, this approach is rigorous and if the titration cell dimensions could be decreased in the future, then lower amounts of materials would be required.

Surface Plasmon Resonance to Measure the Kinetics of Binding and the K_d

Surface plasmon resonance (SPR) is a technique for measuring the association and dissociation kinetics of ligands (analytes) with a receptor. In SPR, the association of the analyte and receptor with one or the other immobilized on a sensor chip induces a change in the refractive index of the layer in contact with a gold film (Figure 27.5). This is measured as a change in the refractive index at the surface layer and is recorded as the SPR signal or resonance units (RU). Such measurements are often conducted in an instrument made by Biacore Life Sciences called the BIAcore™. Binding is measured in real time (Figure 27.5), and information about the association and dissociation kinetics can thus be obtained, which in turn can be used to obtain K_a and K_d from Equations 2 and 3.

A variety of chemistries are available for the coupling of the ligand or receptor to the surface of the chip, including reaction with amines, thiols, aldehydes, and noncovalent biotin capture. In some approaches, a glycoprotein ligand for a GBP is immobilized and the binding of the GBP is measured directly. It is also possible to degrade the immobilized glycoprotein ligand on the chip sequentially by passing over solutions containing exoglycosidases and reexamining at each step the binding to different GBPs, thereby obtaining structural information about the ligand. The immobilized ligand is usually quite stable and can be used repeatedly for hundreds of runs during a period of months.

The advantages of this approach are that (1) affinities in the range from millimolar to picomolar can be measured; (2)

complete measurements of k_1 and k_2 are routine (see Equations 2 and 3) and calculations of K_d are straightforward; (3) for immobilization of a molecule using amine coupling, only 1–5 μg is normally sufficient; (4) typically, the concentration range of analyte is 0.1–100 × K_d and the typical volumes needed are in the range of 50–150 μl; and (5) measurements are extremely rapid and complete experimental results can be obtained within a few days. The drawbacks of this approach are that (1) analytes must have sufficient mass to cause a significant change in SPR upon binding (thus, the glycan is usually immobilized instead of the protein), (2) coupling of free glycans to the chip surface is inefficient and thus neoglycoproteins or some other type of large conjugate must be immobilized, and (3) there may some nonhomogeneity in conditions on the BIAcore™ because of mass transport effects, which could affect the dissociation rate and thus provide an inaccurate K_d measurement.

Fluorescence Polarization for Measuring K_d

A relatively new technique for measuring the binding constant of glycans to GBPs is fluorescence polarization. This approach is based on the reduced rotational motion of a relatively small glycan when it is bound to a relatively large protein compared to the free glycan's rotation in solution. The technique utilizes a glycan containing a fluorophore and a filter to select molecules oriented close to the plane of incident polarized light. Light absorbed by the fluorophore is emitted as fluorescence, but the angle of the emission relative to the incident light is depolarized by rotation of the molecule in solution. In practice, a fluorescently labeled glycan is incubated with increasing concentrations of a GBP and the fluorescence depolarization is measured. In the absence of the GBP, the fluorescently labeled glycan tumbles randomly and the degree of polarization is low. However, if the fluorescently labeled glycan binds to the GBP, its rate of rotation is diminished and the polarization remains high. By this approach, one can measure directly the K_d of the interaction as a function of the GBP concentration.

The advantages of this technique are that (1) the technique is a homogeneous assay and provides direct measurements of the K_d in solution without derivatization of the GBP, (2) it is relatively simple and can provide rapid measurements of many compounds using microtiter plate-based approaches, (3) it utilizes relatively small amounts of glycan, (4) the concentrations of all the molecules are known, (5) it avoids complications of multivalent interactions because the glycans are monovalent and free in solution, and (6) it is amenable to inhibition by competitive agents and can be used to determine relative potency of compounds as inhibitors of GBPs. In the latter approach, a single fluorescently labeled glycan is mixed with the GBP, increasing concentrations of inhibitor glycans (which are not fluorescently labeled) are added, and the inhibition of binding is measured. Because the interactions are simple single-site competition, it is possible to use the concentration that causes 50% inhibition (IC_{50}) to derive the K_d for the inhibitor. Some of the disadvantages are that (1) the technique is limited to small molecules (≤2000 daltons), (2) it requires fluorescence derivatization of the glycan (the fluorophore may alter the properties of the glycan), and (3) preparation of the glycan and chemical derivatization may be tedious and require large amounts of glycans. However, once the fluorescently labeled glycans are generated, there is usually enough for thousands of assays.

Nonequilibrium Methods

ELISA to Measure Specificity and Relative Binding Affinity of Ligands

The conventional enzyme-linked immunosorbent assay (ELISA) has been adapted for studying glycans and GBPs in a variety of formats. Of course, many glycans are antigens and antibodies to them can be analyzed in the conventional ELISA format. Some of the earliest approaches used biotinylated bacterial polysaccharides captured on streptavidin-coated microtiter plates to measure interactions of antibodies to the polysaccharides. In most types of ELISAs used in

glycobiology, either an antibody or a GBP of interest is immobilized and the binding of a glycan to the protein is measured, or the reagents are reversed. In either approach, the glycans are conjugated in some way, such as to biotin or to another protein with an attached reporter group (e.g., a fluorescent moiety or an enzyme such as peroxidase).

Competition ELISA-type assays have also recently been developed to probe the binding site of a GBP or an antibody. In this approach, a glycan is coupled to a carrier protein (the target) and its binding to an immobilized GBP is detected directly. Competitive glycans are added to the wells and their competition for the GBP is measured as a function of concentration to obtain an IC_{50}. The major advantages of this approach are that (1) it is relatively easy, (2) it has high-throughput capability and can be used in an automated fashion by robotic handling, (3) it can provide relative K_d if the GBP concentration is varied appropriately over a large range and binding is saturable, and (4) it has the capacity to define the relatively binding activity of a panel of glycoconjugates. The major disadvantages are that (1) it does not provide direct information about affinity constants or other thermodynamic parameters, (2) it can require relatively high amounts of GBP and glycans if used as a general screening array, and (3) it usually requires chemical derivatization of glycans or GBPs.

Glycan Microarrays to Measure Specificity

Glycan microarrays are an extension of both ELISA-type formats and modern DNA and protein microarray technology. In the microarray, glycans are captured, usually covalently, through reaction with *N*-hydroxysuccinimide (NHS)- or epoxide-containing surfaces on a glass slide. Glycans are printed, much like DNA is printed for DNA microarrays, using contact printers or piezoelectric (noncontact) printing (Figure 27.6). Usually, a few nanoliters of solutions containing glycans in concentrations of 1–100 μM are deposited by a robotic printer on the glass surface in 5–15 μm-diameter spots. Slides are incubated for several hours to allow the chemical reactions to covalently fix the samples on the slides. In most cases, glycans are prepared to contain reactive primary amines at their reducing termini, although other chemical coupling methods are available. These microarrays are then overlaid with a buffer containing the GBP and incubated for several hours to allow maximal binding to occur. The slides are washed to remove unbound GBP and then analyzed. Analyses involve fluorescence detection, which means that either the GBP has to be directly fluorescently labeled or fluorescently tagged antibody to the GBP must be used.

The chief feature of the successful microarrays are the variety of glycans they contain and the clustered and high-density presentation of glycans that promotes binding of even relatively low affinity GBPs. Thus, the density of the ligand should be taken into account when interpreting the results. A publicly available microarray and data on many different GBPs analyzed on the microarray are available from the Consortium for Functional **Glycomics**. In a typical successful analysis, several glycans may be bound by the GBP and appear as intensely fluorescent spots against the background. For example, in Figure 27.7, the binding of the fluorescently labeled influenza virus (strain A/Oklahoma/323/03) to the microarray is shown. Six major glycans, each containing terminal α2-6-linked sialic acid, are clearly bound above background, and hundreds of other glycans on the microarray did not bind appreciably. The data are visually imaged on a scanner and then graphically represented. If desired, the GBP can then be tested for its binding to the identified candidates by other methods to define the K_d, such as titration microcalorimetry or fluorescence polarization, as discussed above. The use of microarrays in characterizing GBPs is a central component of functional **glycomics** (see Chapter 48).

Agglutination

In this approach, one measures the ability of a soluble glycan to block the ability of a multivalent GBP to agglutinate

cells expressing glycans recognized by the GBP. The concentration of the soluble glycan that provides 50% inhibition of agglutination is taken as the inhibitory concentration (IC_{50}). Such approaches have been used for many years in studies on lectin agglutination of cells and were useful in elucidating the nature of the human blood group substances. If a sufficiently large panel of soluble glycans is used, then the relative efficacies of each of these can be measured to help define the specificity of the GBP. A major advantage of this technique is that it does not require tagging of the glycans. Furthermore, polystyrene or dextran beads modified with discreet glycans can be used in lieu of cells. In this case, the glycans on the agglutinating particle are better defined. Usually, the IC_{50} does not relate directly to the binding affinity, since inhibition is being measured. The actual binding affinity must be defined by other techniques described earlier in this chapter.

Precipitation

The interaction of a multivalent GBP or antibody with a multivalent ligand allows for the formation of cross-linked complexes in solution. In many cases, these complexes are insoluble and can be identified as precipitates. In this technique, a fixed amount of GBP or antibody is titrated with a glycoprotein or a glycan bound by the protein of interest. At a precise ratio of ligand to receptor, a precipitate is formed. Such precipitation may be highly specific and reflect the affinity constant of the ligand for the receptor. The amount of protein or ligand in the precipitate can be measured directly by chemical means, using assays for glycans or proteins. The technique of precipitation is still useful for studying potentially multivalent ligands, and it has been used recently to demonstrate that each branch of terminally galactosylated complex-type di-, tri-, and tetra-antennary N-glycans is independently recognized by galactose-binding lectins. Another precipitation approach takes advantage of the fact that a complex between a GBP and a glycan can be “salted out” or precipitated by ammonium sulfate. A variation of this approach was used in early studies on the characterization of the hepatocyte Gal/GalNAc receptor (asialoglycoprotein receptor), in which the ligand (in this case ^{125}I-labeled asialoorosomucoid) was incubated with a preparation of receptor. The sample was treated with an amount of ammonium sulfate capable of precipitating the complex, but not the free unbound ligand. The precipitated complex was captured as a precipitate on a filter and the amount of ligand in the complex was directly determined by γcounting.

Electrophoresis

In this approach, a glycoprotein (or ligand) is mixed with a GBP or antibody and the mixture is electrophoretically separated in polyacrylamide. For glycosaminoglycans, this technique is termed affinity coelectrophoresis (ACE) (see Chapter 16). This method is particularly useful in defining the apparent K_d of the interaction and allows for identification of subpopulations of glycosaminoglycans that differentially interact with the GBP. In another method, termed crossed affinity immunoelectrophoresis, a second step of electrophoresis is conducted in the second perpendicular dimension across an agarose gel that contains precipitating monospecific antibody to the glycoprotein or ligand. The gel is then stained with Coomassie brilliant blue and an immunoelectrophoretogram is obtained. Glycoproteins not interacting with the GBP or antibody have faster mobility than the complex. The amount of glycoprotein or ligand is determined by the area under the curves obtained in the second dimensional analysis. This method is useful for studying glycoforms of proteins and has been particularly valuable in analyzing glycoforms of α1-acid glycoprotein (an acute-phase glycoprotein) in serum and changes in its α1-3-fucosylation.

Expression of cDNAs for Ligands and Receptors

A very indirect approach to studying glycan–protein interactions is to express the cDNA encoding either a

glycosyltransferase or a GBP in an animal or bacterial cell. The adhesion of the modified cell to a GBP or antibody is then measured and taken to reflect the binding of the GBP or antibody to the new glycans (neoglycans) on the cell surface. For example, this approach has been highly valuable in studying selectin ligands and helped to lead to the identification of sialyl Lewis X and sialyl Lewis A as important recognition determinants for selectins and the expression cloning of the cDNA encoding the PSGL-1 (see Chapter 31). The expression of selectins or I-type lectins on the cell surface of transfected cells has been helpful in evaluating the specific role of GBPs in cell adhesion under physiological flow conditions (see Chapters 31 and 32).

MODELING GLYCAN–PROTEIN INTERACTIONS

Computational methods have been applied to studies of glycan conformation. Glycans can assume numerous conformations in solution because of flexibility about glycosidic bonds and internal fluctuations in the rings (see Chapter 2). These conformations are dynamic and change rapidly on a nanosecond–microsecond timescale. Thus, the static structures depicted in diagrams represent average structures, which oftentimes may not be representative of the predominant structure that exists in solution. Molecular dynamics simulations can predict these conformational states and determine their relative probability. The relationships between sequence and conformational behavior are highly variable and not all parameters can be fully taken into account (e.g., the role of water, pH, and electrostatic charge). The picture that emerges from these studies is one of highly dynamic molecules, in which torsional rotations about glycosidic bonds are accompanied by internal fluctuations, allowing the molecule to assume stable conformations.

Molecular modeling also can be used to examine glycan conformations when bound to a GBP. These studies usually take into account other experimental data obtained by NMR experiments, X-ray crystallography, and cross-linking studies. Molecular modeling can also be used to predict the relative binding energies for closely related ligands, thus aiding the selection and development of potential antagonists. As computer capacity continues to improve and new algorithms are created to take into account all atoms in the system, modeling will most likely become more commonplace.

GLYCAN–GLYCAN INTERACTIONS

In addition to GBP–glycan interactions, a growing body of evidence suggests that glycan–glycan interactions also may be of importance, especially in the context of cell–cell contacts. Complementary binding between glycans is often considered nonspecific, but biologically relevant high affinities might be achieved by the multivalent interactions made possible by the high density of glycoconjugates on the plasma membrane. One of the best studied biological systems is the sponge, where species-specific cell–cell recognition can occur through multivalent interactions between high-molecular-mass (~200 kD) glycans. The composition of these glycans from different individuals of the same species does not vary, which allows species-specific recognition between cells. In contrast, the composition of glycans on cells from different species varies and prevents their aggregation. Other examples of glycan–glycan interactions include Ca^{++}-dependent cell–cell interactions mediated by glycosphingolipids. For example, interactions between glycosphingolipids expressing the $Lewis^x$ blood group determinant are thought to facilitate compaction of mouse embryonic cells in the 8–32 cell stage and formation of synapses across myelin sheets in the nervous system. Sequestration of glycosphingolipids in microdomains (lipid rafts) of the plasma membrane might provide points of multivalent contact between cells. The molecular forces active between glycans are not well-characterized but presumably involve the same interactions that facilitate protein–glycan interactions, such as van der Waals contacts, hydrogen bonding, and hydrophobic interactions described in previous sections. The requirement for Ca^{++} ions in many of these systems may indicate their role in the clustering of glycans or in conformational changes

in the sugar chains to provide optimal interactions. Further studies are needed to understand the importance and frequency of glycan–glycan interactions in various biological processes.

FURTHER READING

1. Rini JM. Lectin structure. Annu Rev Biophys Biomol Struct. 1995;24:551–577. [PubMed: 7663127]
2. Weis WI, Drickamer K. Structural basis of lectin–carbohydrate recognition. Annu Rev Biochem. 1996;65:441–473. [PubMed: 8811186]
3. von der Lieth C, Siebert H, Kozar T, Burchert M, Frank M, Gilleron M, Kaltner H, Kayser G, Tajkhorshid E, Bovin NV, Vliegenthart JF, Gabius H. Lectin ligands: New insights into their conformations and their dynamic behavior and the discovery of conformer selection by lectins. Acta Anat. 1998;161:91–109. [PubMed: 9780353]
4. Woods RJ. Computational carbohydrate chemistry: What theoretical methods can tell us. Glycoconj J. 1998;15:209–216. [PMC free article: PMC4201040] [PubMed: 9579797]
5. Fred Brewer C. Binding and cross-linking properties of galectins. Biochim. Biophys. Acta. 2002;1572:255–262. [PubMed: 12223273]
6. Duverger E, Frison N, Roche AC, Monsigny M. Carbohydrate-lectin interactions assessed by surface plasmon resonance. Biochimie. 2003;85:167–179. [PubMed: 12765786]
7. Hirabayashi J. Oligosaccharide microarrays for **glycomics**. Trends Biotechnol. 2003;21:141–143. [PubMed: 12679056]
8. Sorme P, Kahl-Knutson B, Wellmar U, Nilsson UJ, Leffler H. Fluorescence polarization to study galectin-ligand interactions. Methods Enzymol. 2003;362:504–512. [PubMed: 12968384]
9. Bucior I, Burger MM. Carbohydrate–carbohydrate interactions in cell recognition. Curr Opin Struct Biol. 2004;14:631–637. [PubMed: 15465325]
10. Dam TK, Brewer CF. Multivalent protein–carbohydrate interactions: Isothermal titration microcalorimetry studies. Methods Enzymol. 2004;379:107–128. [PubMed: 15051354]
11. Feizi T, Chai W. Oligosaccharide microarrays to decipher the glyco code. Nat Rev Mol Cell Biol. 2004;5:582–588. [PubMed: 15232576]
12. Neumann D, Lehr CM, Lenhof HP, Kohlbacher O. Computational modeling of the sugar–lectin interaction. Adv Drug Deliv Rev. 2004;56:437–457. [PubMed: 14969752]
13. Ratner DM, Adams EW, Su J, O'Keefe BR, Mrksich M, Seeberger PH. Probing protein–carbohydrate interactions with microarrays of synthetic oligosaccharides. Chembiochem. 2004;5:379–382. [PubMed: 14997532]
14. Homans SW. Probing the binding entropy of ligand–protein interactions by NMR. Chembiochem. 2005;6:1585–1591. [PubMed: 16038002]
15. Nakamura-Tsuruta S, Uchiyama N, Hirabayashi J. High-throughput analysis of lectin–oligo-saccharide interactions by automated frontal affinity chromatography. Methods Enzymol. 2006;415:311–325. [PubMed: 17116482]
16. Paulson JC, Blixt O, Collins BE. Sweet spots in functional **glycomics**. Nat Chem Biol. 2006;2:238–248. [PubMed: 16619023]
17. Smith DF, Cummings RD. Deciphering lectin ligands through glycan arrays. In: Vasta GR, Ahmed H, editors. Animal lectins: A functional view. CRC Press; Boca Raton, Florida: 2008. (In press)

Figures

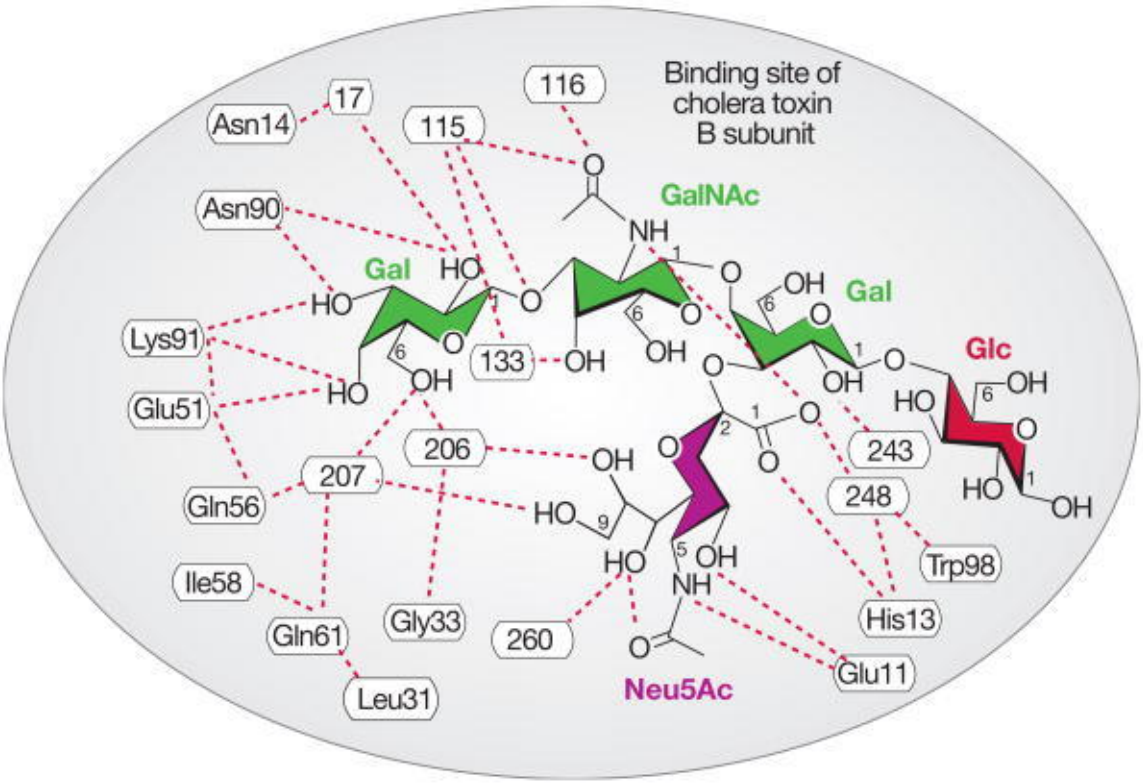


FIGURE 27.1

Simplified structure of the GM1 pentasaccharide bound to the B subunit of cholera toxin. Identification of the amino acid residues in cholera toxin that contact the glycan directly or through hydrogen bonds with water. (Redrawn, with permission, from Merritt E.A., Sarfaty S., van den Akker F., et al. 1994. *Protein Sci.* 3: 166–175.)

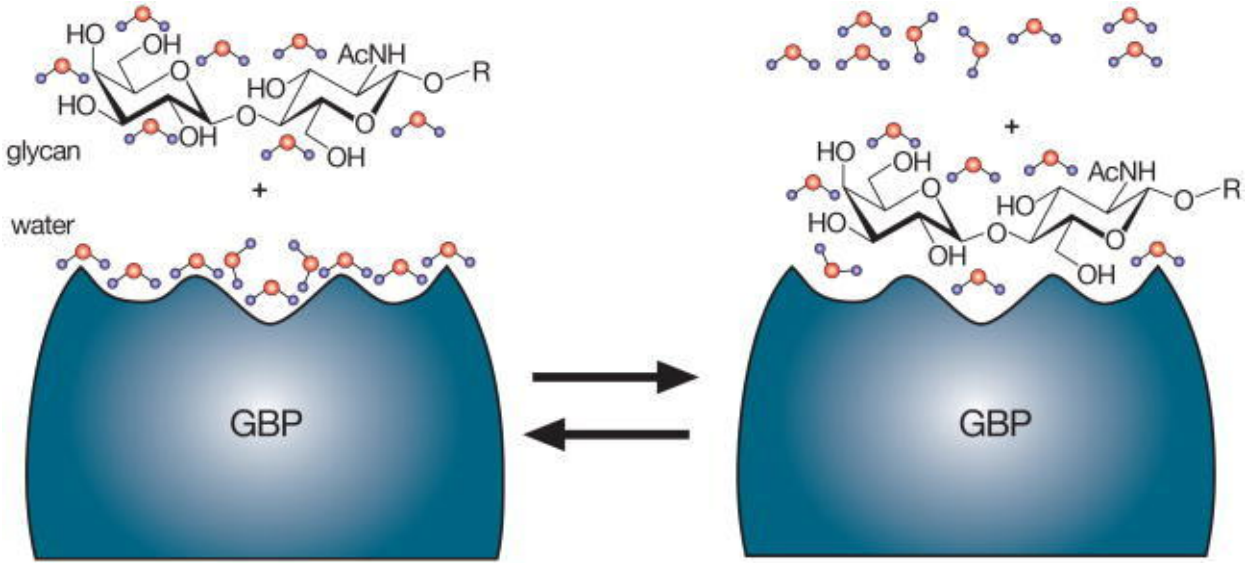


FIGURE 27.2

Schematic diagram of the binding of a glycan to a GBP in water, which results in the displacement of water.

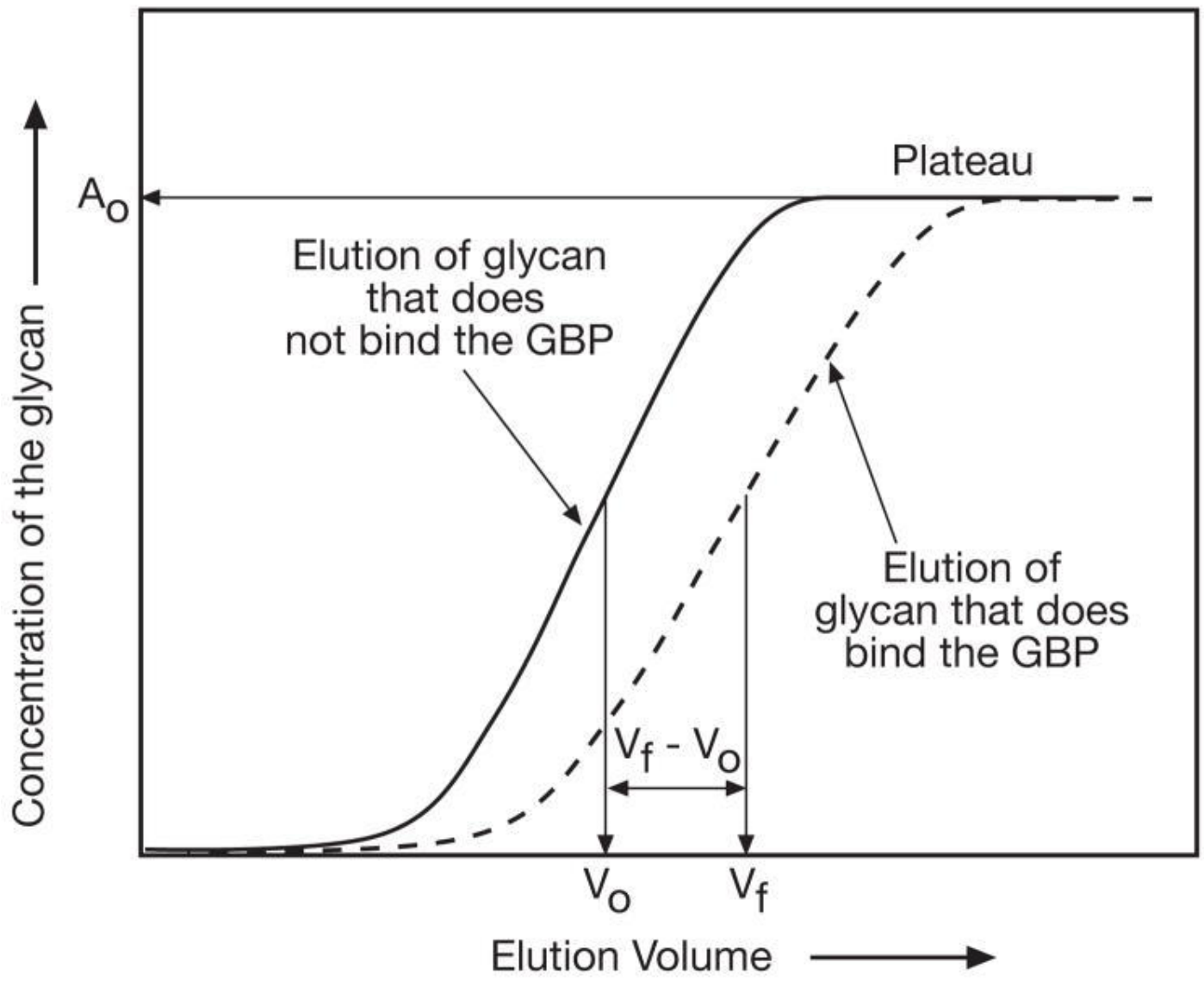


FIGURE 27.3

Example of frontal affinity chromatography, where different concentrations of a glycan are applied to a column of immobilized GBP. The profile depicts the elution of one glycan that binds the GBP and the elution of another glycan that does not bind the GBP. The V_o (void volume) and V_f are determined as shown from the elution volume. High V_f values indicate higher affinity.

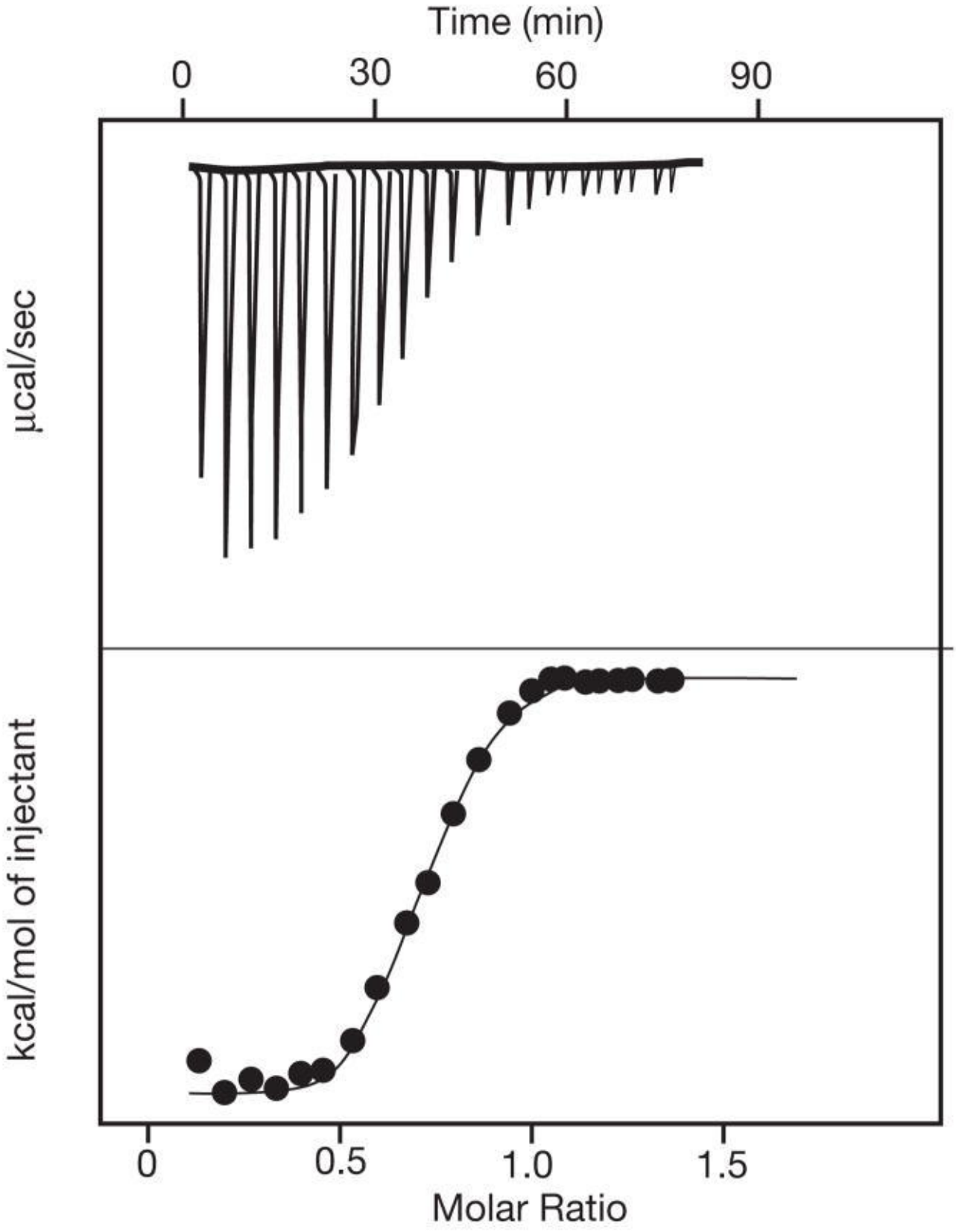


FIGURE 27.4

Example of titration microcalorimetry. Increasing amounts of a glycan are injected to a fixed amount of GBP in a

cell, and the heat produced upon binding is measured as μcal/sec. The total kcal/mole of injected glycan relative to the molar ratio is plotted. These data can be used to define directly the thermodynamic parameters of binding and calculate the K_d of interaction between the glycan and the GBP.

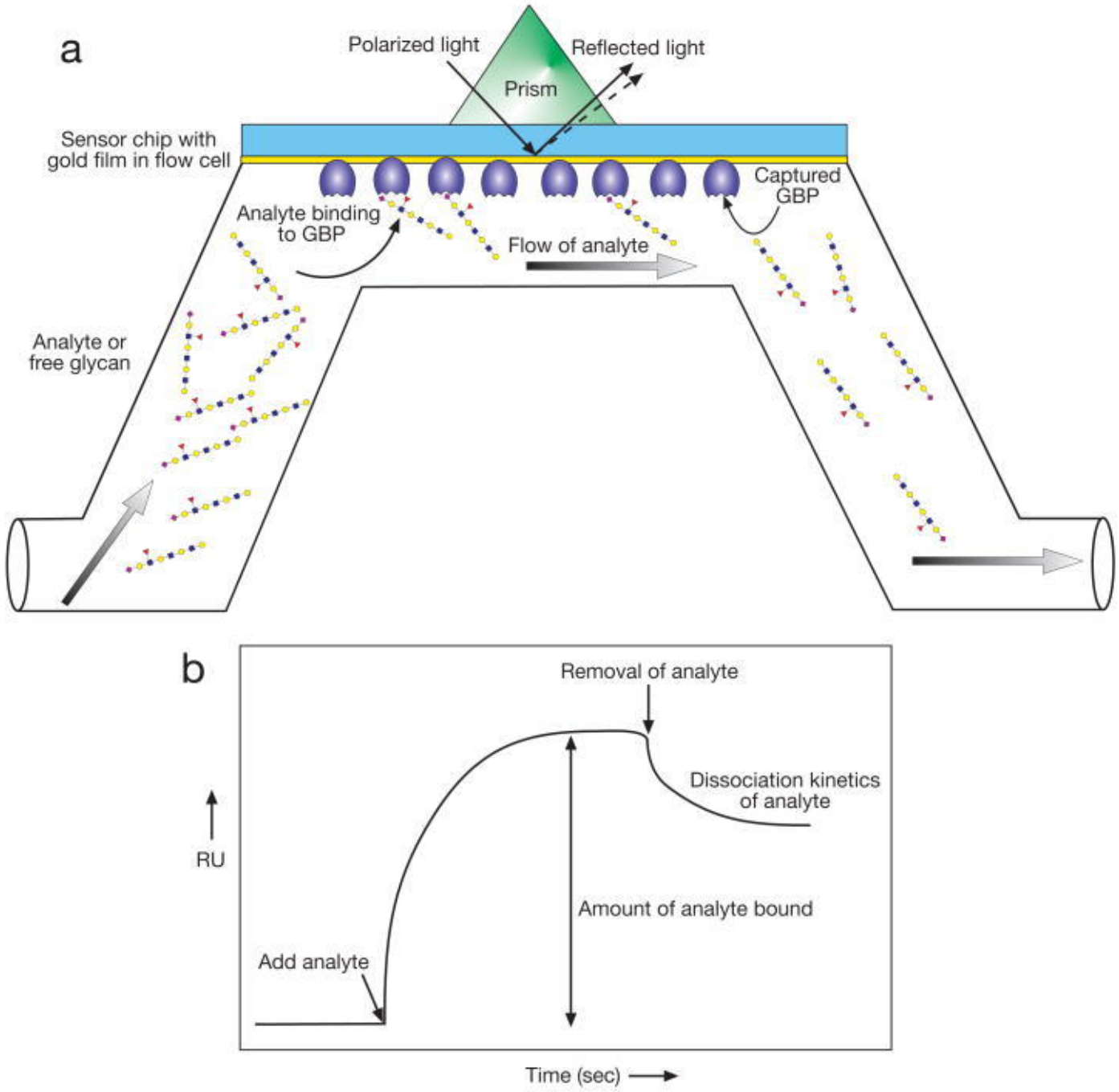


FIGURE 27.5

Example of surface plasmon resonance (SPR). (*a*) In SPR, the reflected light is measured and is altered in response to binding of the analyte in the flow cell to the immobilized GBP. (*b*) An example of a sensorgram showing the binding of the analyte to the ligand and the kinetics of binding and dissociation. RU indicates resonance units.

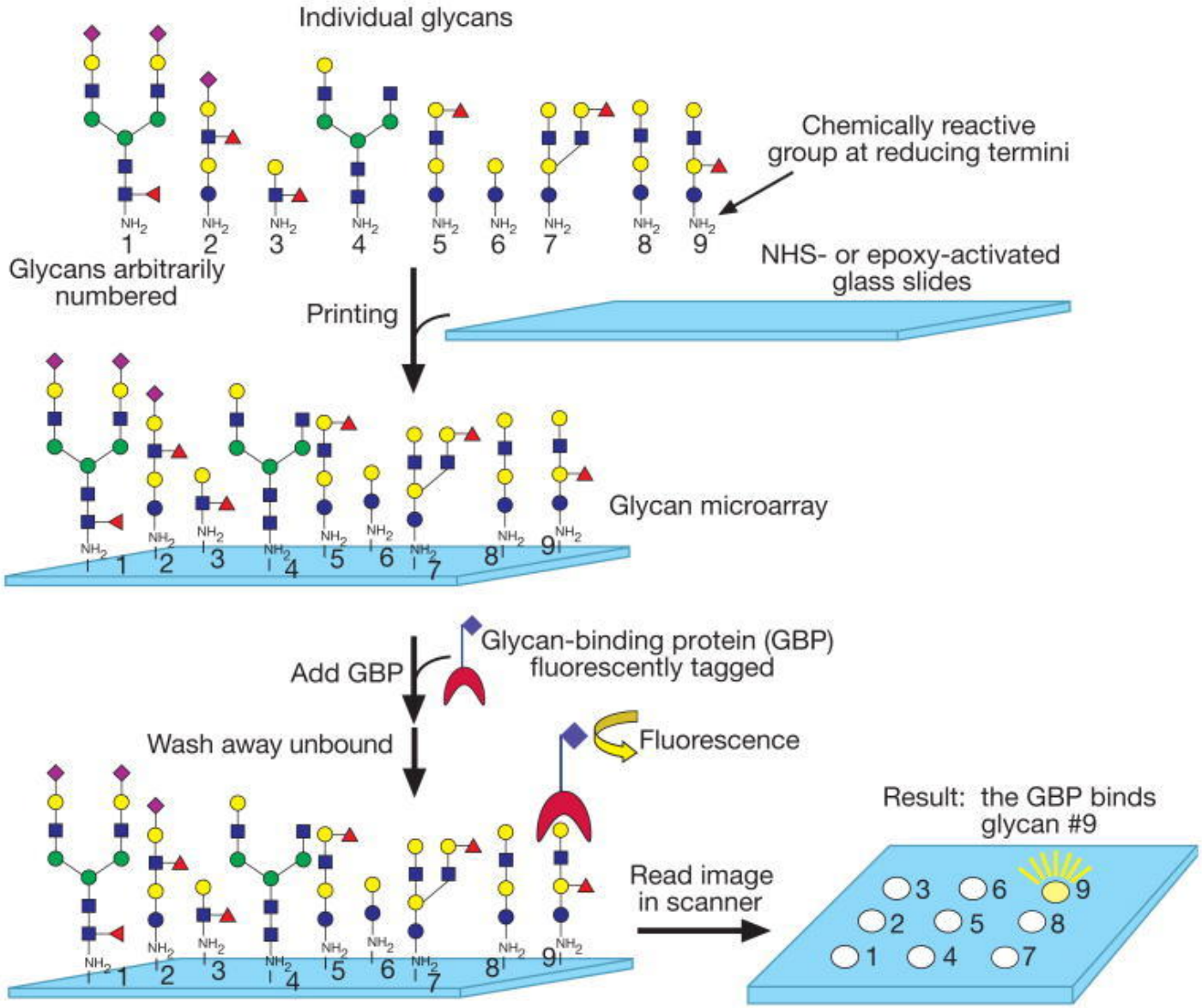
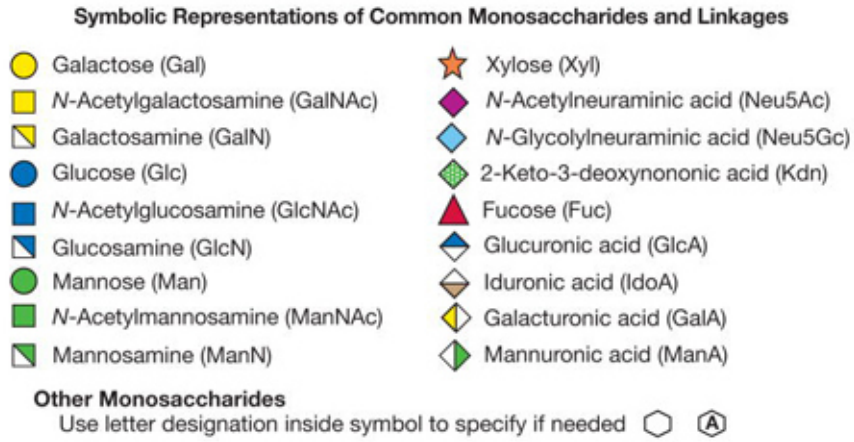


FIGURE 27.6

Preparation of covalent glycan microarrays printed on *N*-hydroxysuccinimide (NHS)- or epoxide-activated glass slides. In this example, the glycans have a free amine at the reducing end and are coupled to the glass slide. After washing, the slide is “interrogated” with a GBP. After washing away the unbound GBP, binding is detected by fluorescence. The GBP may be directly fluorescently labeled or detected using fluorescently labeled secondary reagents. The slides are read in a fluorescence scanner.

Symbol Key:


Symbolic Representations of Common Monosaccharides and Linkages
Galactose (Gal)
N-Acetylgalactosamine (GalNAc)
Galactosamine (GalN)
Glucose (Glc)
N-Acetylglucosamine (GlcNAc)
Glucosamine (GlcN)
Mannose (Man)
N-Acetylmannosamine (ManNAc)
Mannosamine (ManN)
Xylose (Xyl)
N-Acetylneuraminic acid (Neu5Ac)
N-Glycolylneuraminic acid (Neu5Gc)
2-Keto-3-deoxynononic acid (Kdn)
Fucose (Fuc)
Glucuronic acid (GlcA)
Iduronic acid (IdoA)
Galacturonic acid (GalA)
Mannuronic acid (ManA)
Other Monosaccharides
Use letter designation inside symbol to specify if needed
A

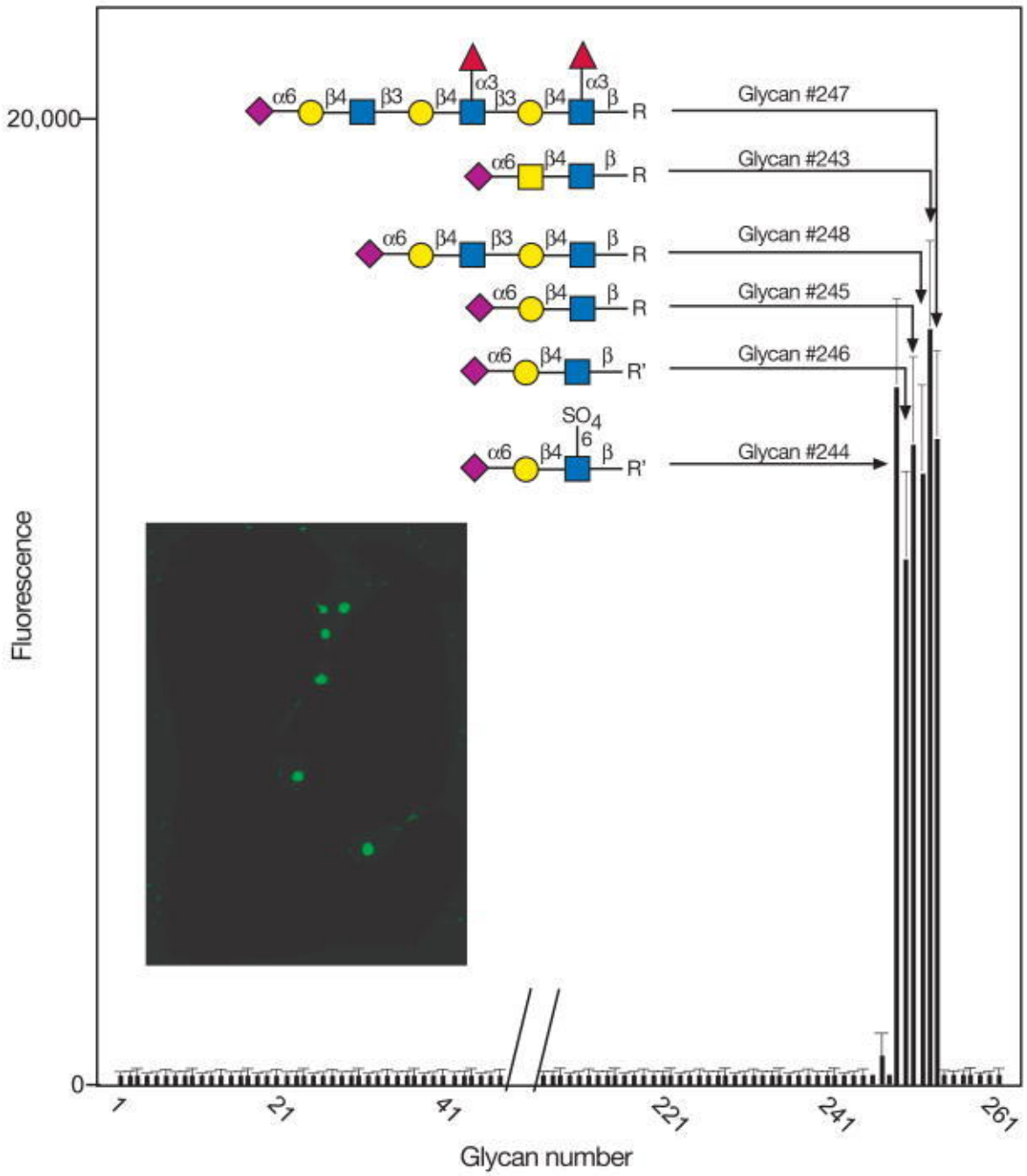
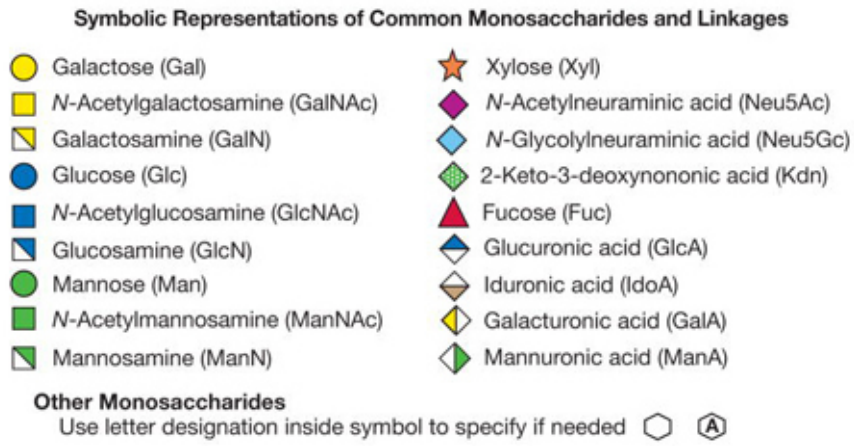


FIGURE 27.7

Binding of a GBP to a covalent glycan microarray. In this example, the fluorescently labeled influenza virus (strain A/Oklahoma/323/03) was labeled with Alexa-488 and applied to the glass slide. The bound virus was detected by fluorescence using a fluorescence scanner. (*Inset*) Type of data obtained for a subarray of the whole slide, where six major fluorescent spots are seen, which represent the six glycans bound well by the GBP. The data are shown as a histogram, where the fluorescent units are plotted versus the glycan number on the microarray. Data are averaged from several replicated patterns of binding in multiple subarrays on the same slide. The structures of the sialylated *N*-acetyllactosamines recognized by this GBP are indicated. (Based on Kumari K., Gulati S., Smith D.F., et al. 2007. *Virology J.* 4: 42, and available at the website of the Consortium for Functional

Glycomics.)

Symbol Key:



Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California.

Bookshelf ID: NBK1950 PMID: 20301269