Chemical tools for functional studies of glycans

Sungjin Park, Myung-Ryul Lee and Injae Shin*

Received 11th March 2008
First published as an Advance Article on the web 9th June 2008
DOI: 10.1039/b713011m

In recent years a variety of chemical approaches have been developed for elucidating the molecular basis of biological processes in which glycans participate. The chemical technologies uncovered have greatly influenced the progress of glycomics research programs. This *tutorial review* highlights recent advances in chemical tools which have been developed and their applications in studies aimed at gaining a better understanding of the roles that glycans play in biological processes.

1. Introduction

Glycans (oligosaccharides and polysaccharides) are a large group of biomolecules with diverse structures that are found inside or on the surface of cells. These substances are present mainly in the form of glycoconjugates, including glycoproteins (glycans linked to asparagine or serine/threonine residues of polypeptides), proteoglycans (glycosaminoglycans linked to proteins) and glycosphingolipids (glycans linked to ceramides). Carbohydrates are also key components of the glycophospholipid anchors which attach proteins to the cell membrane. Glycans affect the activities of the proteins or lipids to which they are attached by modulating their functions. Oligosaccharide groups present in glycoproteins impact the folding and conformational stability of the polypeptide scaffolds. Cell surface glycans serve as ligands for glycan-binding proteins that mediate cell trafficking, adhesion and signaling.[1] In addition, carbohydrate-based biomolecular interactions also play key roles in pathological processes, including tumor metastasis, viral and bacterial infection, and inflammation. Importantly, variations in the sugar composition of glycoconjugates result in certain disease states such as cancer and inflammation. As a result of these properties, glycans serve as ideal targets for the development of agents for clinical diagnostics and therapeutics. Consequently, the biological and biomedical significance of glycans makes studies of their functions one of the most important fields of research in the post-genomic era.

Proteins are produced by template-dependent biosynthetic pathways. In contrast, glycans are biosynthesized in a template-independent manner *via* stepwise processes promoted by glycosyltransferases and glycosidases in the endoplasmic reticulum (ER) and Golgi apparatus (Fig. 1). Thus, technologies that have been used successfully to study protein functions,

Department of Chemistry, Yonsei University, Seoul, 120-749, Korea. E-mail: injae@yonsei.ac.kr

Sungjin Park received her BS degree in Chemistry from the Hankuk University of Foreign Studies in 1998. Her MS degree was awarded in 2001 for studies in carbohydrate chemistry performed at the Yonsei University under the guidance of Professor Injae Shin. She began her PhD with Professor Shin in 2005. Her research interests include the synthesis of glycoconjugates, and the fabrication of carbohydrate microarrays and their applications to the field of glycomics.

Professor Injae Shin received his BS (1985) and MS (1987) degrees in Chemistry from Seoul National University. His PhD studies in Chemistry were carried out at University of Minnesota with Professor Hung-wen Liu (1991–1995). He then moved to University of California-Berkeley where he worked with Professor Peter G. Schultz as a postdoctoral fellow during the period 1995–1998. He started his independent career as Assistant Professor of Chemistry at Yonsei University in 1998 where he became Associate Professor in 2001 and Professor in 2006. His research interests are in the area of chemical biology, including functional studies of glycans using chemical tools, especially carbohydrate microarrays. His research group is also involved in the development of small molecules that modulate protein functions and the construction of target-oriented drug delivery systems. He serves on the editorial advisory boards of ChemMedChem and Curr. Chem. Biol.

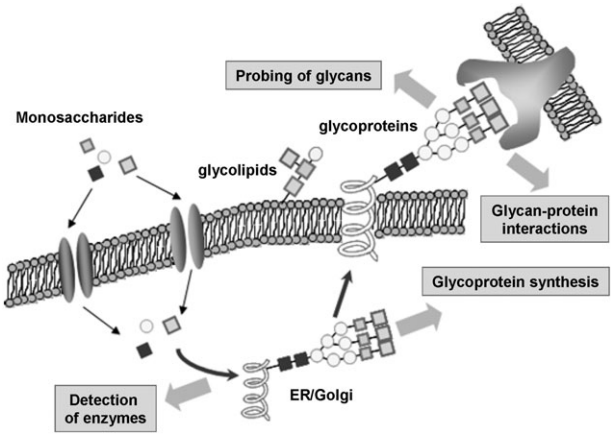

Fig. 1 Biosynthesis of glycoproteins and glycolipids in cells and functional studies of glycans using the chemical tools described in this article.

such as rapid sequencing, site-directed mutagenesis, gene disruption, and *in vivo* and *in vitro* synthesis, are not readily applied to the functional assignment of glycans. Moreover, the micro-heterogeneity in glycan moieties makes the analysis and understanding of their biological roles more difficult. As a result, progress made in understanding the specific roles of glycans and relationships between glycan structure and function has been slow compared with the advances made with proteins.

As efforts made to understand the molecular basis of biological processes in which glycans participate, numerous chemical approaches have been recently developed.[2] As will be demonstrated in this review, chemical methods hold great potential in glycobiology studies. One of the most important chemical tools is associated with the synthesis of complex oligosaccharides. Since synthetic approaches to oligosaccharides have been reviewed extensively,[3] this review will focus on other recent advances made in the development of chemical tools for gaining a better understanding of the molecular basis of the function of glycans in biological processes. The chemical approaches reviewed here include chemical methods for probing carbohydrates, the detection of proteins using synthetic carbohydrate probes, the construction of homogeneous glycoproteins and carbohydrate microarrays for the rapid analysis of carbohydrate-recognition events (Fig. 1).

2. Chemical methods for probing carbohydrates

By way of biomolecular interactions, glycans are involved in various physiological and pathological processes. Importantly, changes in glycosylation of proteins and lipids often cause malignant transformations. As a result, the ability to uncover alterations in glycosylation of glycoconjugates enables a greater understanding of biological processes and disease states that are correlated with glycans. A genetic approach to disrupting genes encoding for glycosyltransferases and glycosidases can be used to elucidate the functional effects of specific glycans. However, this approach has limitations arising from the facts that carbohydrate-processing enzymes are redundant and glycan mutations sometimes lead to lethality during development.[4] On the other hand, biochemical methods relying on antibodies or lectins can be employed to investigate the glycosylation of proteins and lipids. However, it is difficult to detect subtle alterations in glycan structures by using these proteins, owing to their intrinsic binding specificities.

To overcome the limitations of genetic and biochemical approaches, a specific sugar-tagging method has been recently exploited to probe glycosylation in cells and living animals. This method includes the metabolic incorporation of a non-native sugar, appended by a bioorthogonal functional group (a chemical handle), into intracellular or cell surface glycans and subsequent chemoselective labeling of the chemical handle with a biophysical or biochemical probe that can be used for visualization and/or proteomic analysis. A prerequisite for this strategy is that (1) the unnatural glycan containing a chemical handle should be a substrate that can be utilized by the cell's own biosynthetic machinery, and (2) the two ligating partners (a chemical handle and a probe) should form a stable adduct in a specific fashion under physiological conditions. This approach has been applied to probe several sugars, such as sialic acid, *O*-*N*-acetylglucosamine (*O*-GlcNAc), mucin-type *O*-linked glycan and fucose, in cells and living animals (Fig. 2).

The first example, using a sugar-tagging method, is probing sialic acid on cell surfaces.[5] This sugar is the most prevalent terminal residue of glycans on glycoconjugates present in higher organisms and its over-expression is correlated with metastatic and malignant phenotypes in tumors.[6] To detect this sugar, ketone-containing sialic acid is incorporated into glycoconjugates in cells by incubating them with peracetylated *N*-levulinoyl mannosamine (Ac_4ManLev) (Fig. 2a).[5] Sugars protected by acetyl groups are normally used in this tagging approach due to their higher efficiency of uptake by cells in comparison to that of free sugars.[7] The acetyl protecting groups of the sugars are removed once inside cells by ester hydrolysis reactions catalyzed by cytosolic esterases. ManLev in cells is converted to the corresponding sialic acid, which is then integrated into sialosides by metabolic biosynthetic pathways. The ketone handle that is exposed in cell surface sialosides is used for selective ligation to hydrazide- or aminnoxy-conjugated probes under physiological conditions (Fig. 3). This methodology has been applied in studies of expression patterns of sialic acids on cell surfaces, targeting of cancers using an MRI technique, and viral-mediated gene delivery after introducing artificial virus receptors.

The ketone handle, although amenable to monitoring cell-surface sugars, has limited utility for the detection of sugars in intracellular glycoconjugates and living organisms because metabolites containing ketone or aldehyde, such as free carbohydrates, pyruvate and oxaloacetate, are abundant. To circumvent this problem, azide functionality can be used as a bioorthogonal handle since it is rarely present in naturally occurring substances and it is nearly inert in cells and living animals. Glycoconjugates possessing azidosialic acid are generated in cells or organism by incubating with a peracetylated form of *N*-azidoacetylmannosamine (Ac_4ManNAz) or by injecting the sugar into mice (Fig. 2a).[8,9] The azide-containing glycoconjugates are then probed by treatment of the cells or tissues isolated from the treated animals with biochemical probe-conjugated phosphine reagents *via* Staudinger ligation

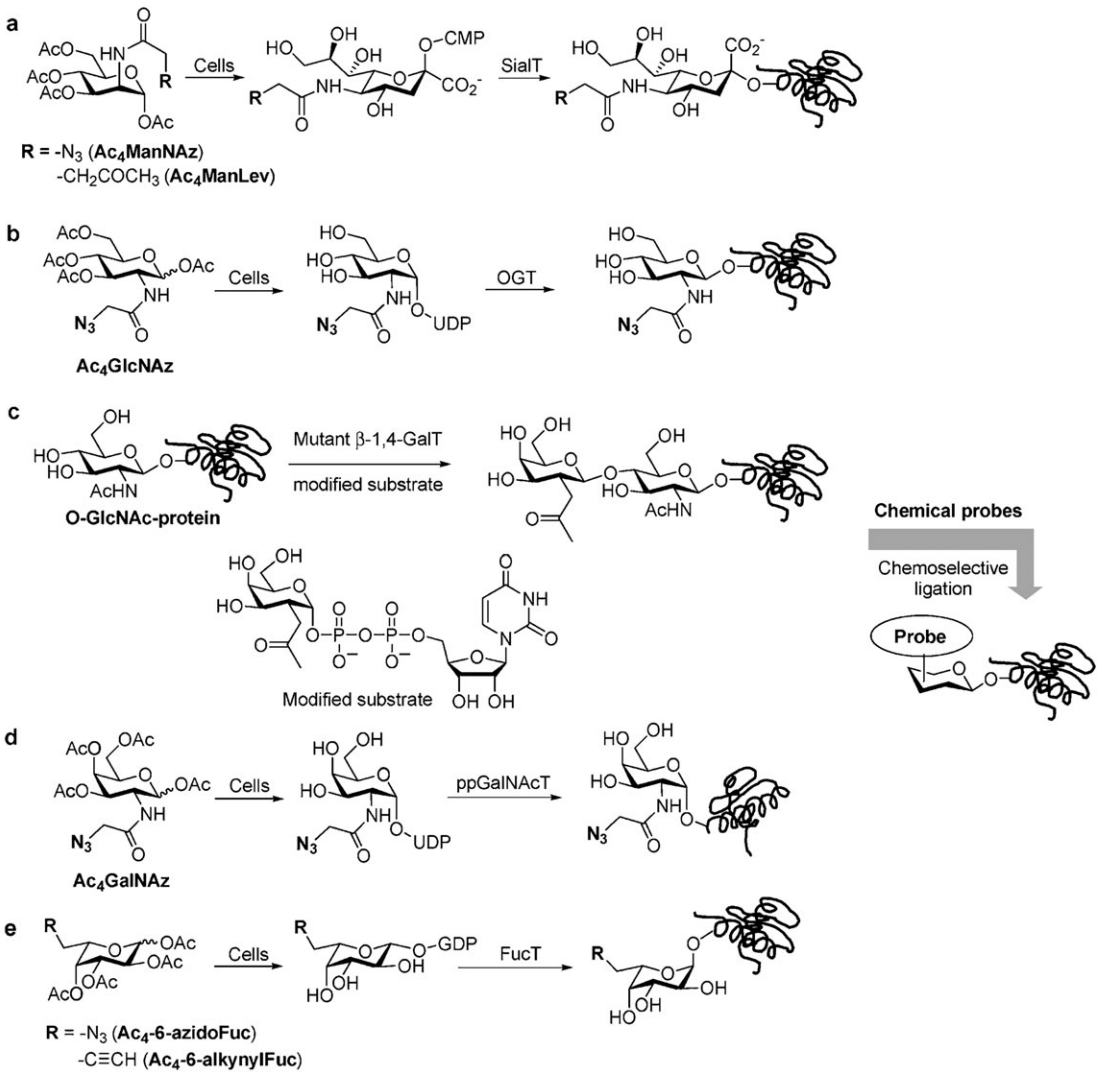

Fig. 2 Sugar-tagging methods to probe glycosides in intracellular and cell surface glycoconjugates (SialT; sialyltransferase, OGT; GlcNAc: polypeptidyltransferase, ppGalNAcT; polypeptide *N*-acetyl-α-galactosaminyltransferase, FucT; fucosyltransferase).

(Fig. 3). It has been shown that the fluorescence signal arising from the azidosugar is higher than that associated with the ketosugar, suggesting that ManNAz is a better substrate for sialic acid biosynthesis than ManLev. This is possibly a result of the fact that the structure of ManNAz is more similar to that of natural ManNAc.

A single GlcNAc residue (*O*-GlcNAc), which is attached to the serine and threonine residues of nuclear and cytoplasmic

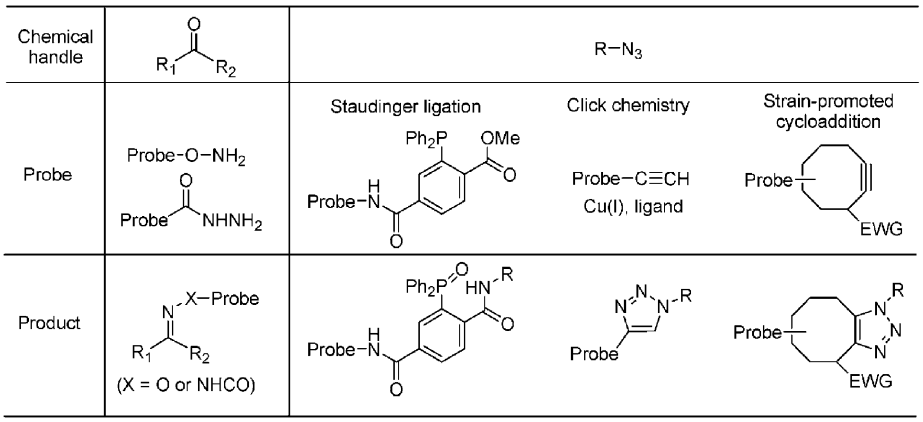

Fig. 3 Chemoselective ligation used to detect modified sugars in proteins, cells and organisms (EWG; electron withdrawing group).

proteins, plays an important role in the regulation of multiple cellular pathways in a variety of organisms. The dysregulation of *O*-GlcNAc modification is known to lead to various disease states including diabetes, cancer and Alzheimer's disease. The sugar-tagging approach has been applied to probe *O*-GlcNAc in cells and tissues and for the proteomic analysis of *O*-GlcNAc-proteins.[10] For these purposes, cells are incubated with peracetylated *N*-azidoacetylglucosamine (Ac_4GlcNAz) to biosynthetically incorporate the azide into *O*-GlcNAc-proteins (Fig. 2b). The azide handle in *O*-GlcNAc-proteins from cell lysates is selectively linked to biochemical probe-conjugated phosphine reagents by Staudinger ligation. This procedure has been used to identify 199 putative *O*-GlcNAc-proteins from HeLa cells.[11]

Methods that employ nonnative GlcNAz have limited applications to the proteomic analysis of *O*-GlcNAc-proteins because the nonnative sugar is not incorporated into all *O*-GlcNAc-proteins in cells. To directly analyze *O*-GlcNAc-proteins, an approach using an engineered β-1,4-galactosyltransferase (β-1,4-GalT) and a ketone-containing modified substrate has been developed.[12] This methodology involves initial galactosylation of *O*-GlcNAc-proteins from cell lysates by treatment with an engineered β-1,4-GalT and a UDP-ketone containing galactose analog (Fig. 2c). The engineered β-1,4-GalT can transfer a ketone-containing galactose analog selectively to the C-4 hydroxyl group of GlcNAc in *O*-GlcNAc-proteins in cell lysates.[13] Subsequently, ketone-containing proteins are labeled with an aminooxy-conjugated biotin. Protein digestion by protease, followed by enrichment of *O*-GlcNAc peptides by avidin affinity chromatography and mass spectrometric analysis of the enriched glycopeptides leads to the rapid identification of the *O*-GlcNAc-proteins and the determination of *O*-GlcNAc modification sites in the glycoproteins.[14]

Mucin-type *O*-linked glycosylation, in which *N*-acetylgalactosamine (GalNAc) is attached to serine or threonine residues of the protein backbone with α-anomeric configuration, is the most popular form of *O*-glycosylation. This type of glycosylation can be assessed by using a sugar-tagging technique in which a peracetylated *N*-azidoacetylgalactosamine (Ac_4GalNAz) is metabolically installed onto mucin-type *O*-linked glycoproteins in cells (Fig. 2d).[15] The glycoproteins containing the azide in cells are detected by labeling with a chemical probe *via* Staudinger ligation. This approach, applied to the rapid profiling of *O*-linked glycoproteins in living animals by injecting Ac_4GalNAz into mice, has demonstrated that azido-labeled glycoproteins appear in various tissues, such as liver, kidney and heart, in serum, and on isolated splenocytes.[16]

Fucose serves as epitopes involved in regulation of protein function and cell–cell interactions. For example, Le^x-displaying glycoproteins play a role in neurogenesis and promote embryonic cell adhesion. Changes in the fucosylation of glycoproteins result in several pathological processes such as inflammation and tumorigenesis. For probing protein fucosylation, 6-azidofucose or 6-alkynylfucose is biosynthetically inserted into glycoproteins in cells by incubating with the respective peracetylated form of 6-azidofucose (Ac_4-6-azidoFuc) or 6-alkynylfucose (Ac_4-6-alkynylFuc) (Fig. 2e).[17] Unlike the other types of azidosugars described above, 6-azidofucose has a relatively high cytotoxicity. However, 6-alkynylfucose displays a greatly reduced toxicity to cells in comparison to its azido counterpart. Cell surface and intracellular fucosides in glycoconjugates are visualized by incubating cells with Ac_4-6-alkynylFuc and subsequent labeling of the alkyne-conjugated fucosides with 3-azido-7-hydroxycoumarin *via* Cu(I)-catalyzed cycloaddition (click chemistry).

An important component using azide-based probes for detecting glycans in cells or animals is the ability to selectively and efficiently label the azide-containing molecules with functionalized chemical probes. In general, labeling of the azide has been achieved by using three methods including click chemistry, Staudinger ligation and Cu(I)-free strain-promoted cycloaddition (Fig. 3). Click chemistry is a very sensitive azide labeling method and consequently it has been widely used for probing enzyme activities in cell lysates and visualizing biomolecules in cells. However, the method is not applicable for studying dynamic processes in living systems because of the need to use a cytotoxic copper catalyst. The Staudinger ligation does not need cytotoxic reagents and thus it is highly biocompatible with cells and living animals. However, this labeling procedure undergoes slow ligation, which causes competitive and undesirable oxidation of the phosphine reagents. Therefore, this ligation process is not well suited to monitor species with low abundance and to follow rapid biological processes. Processes involving strain-promoted cycloaddition with electron-withdrawing groups have been developed for the sensitive and rapid labeling of azides. Unlike conventional click chemistry, this labeling method does not require a copper catalyst.[18] Overall, the sugar tagging approach combined with an efficient labeling technique enables the study of dynamic processes of glycosylation in living systems. Moreover, this approach may be ultimately utilized for diagnostic and therapeutic purposes since changes in glycosylation are observed in various disease states.

3. Detection of proteins using synthetic carbohydrate probes

3.1 Activity-based probes to detect glycosidases

Glycans present in glycoconjugates are mainly assembled in the endoplasmic reticulum and Golgi by glycosyltransferases and glycosidases. In order to obtain global information relating to the biosynthesis and function of glycans, studies of these enzyme activities are of great importance in glycomics research. Synthetic probes (termed activity-based probes) that selectively label enzymes by means of their catalytic activities in complex mixtures have been used as powerful tools for functional profiling of various enzyme families, such as proteases, kinases, phosphatases, histone deacetylases and oxidoreductases.[19] This approach has also been applied to the identification and characterization of glycosidases in cells. The activity-based probes for glycosidases contain (1) a mechanism-based inhibitor moiety that reacts with an active-site nucleophilic residue in the enzyme to form a covalent adduct, and (2) a reporter moiety that allows for rapid detection (*e.g.* a fluorescent dye) and/or affinity purification (*e.g.* biotin) of the labeled target enzyme.

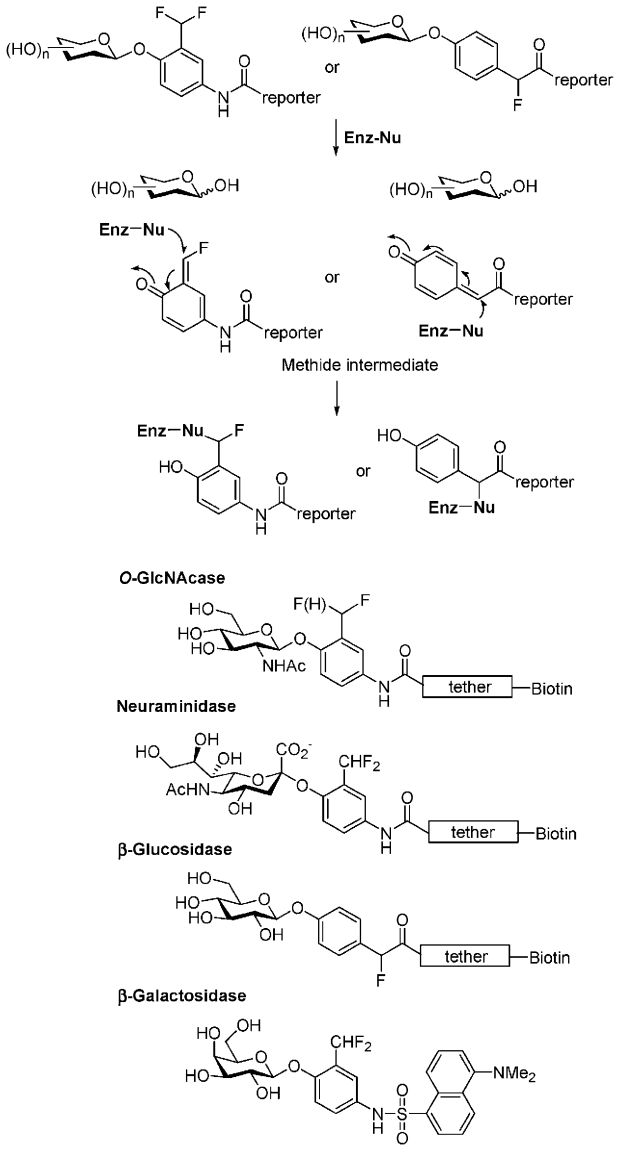

Fig. 4 Activity-based probes to detect glycosidases using fluoromethylated aryl glycosides.

Two types of mechanism-based inhibitors, fluoromethylated aryl glycosides and 2-deoxy-2-fluoroglycosides, have been utilized to capture glycosidases in cells. In the case of a fluoromethylated aryl glycoside probe, cleavage of the glycosidic linkage by glycosidase liberates a reactive methide intermediate by 1,4- or 1,6-elimination (Fig. 4).[20] A nucleophilic residue inside or near the active site attacks the methide intermediate to form a covalent adduct. Since the activity-based probe contains either a fluorescent dye or biotin as a reporter, the enzyme-inhibitor adduct can be readily detected by fluorescence analysis or isolated by affinity column chromatography. This type of the activity-based probe has been employed to detect *O*-*N*-acetylglucosaminidase from rat spleen (*O*-GlcNAcase), β-glucosidase, neuraminidase from influenza viruses and β-galactosidases from several bacteria.[21–24]

Another type of mechanism-based inhibitor is the 2-deoxy-2-fluoroglycoside that contains a good leaving group at the

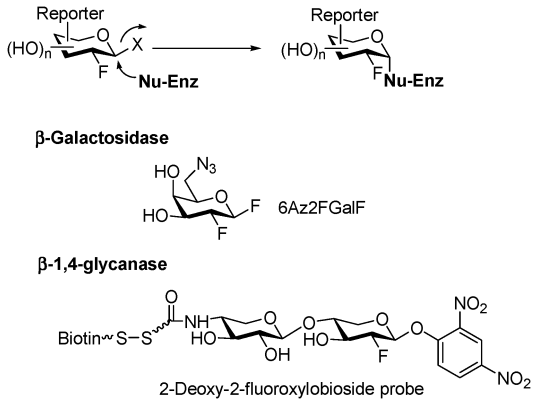

Fig. 5 Activity-based probes to detect glycosidases using 2-deoxy-2-fluorosugars.

anomeric position. This inhibitor can capture retaining glycosidases that catalyze cleavage of the glycosidic linkage through a double-displacement mechanism involving formation and subsequent breakdown of a covalent glycosyl–enzyme intermediate (Fig. 5). The anomeric carbon in a 2-deoxy-2-fluoroglycoside probe is attacked by the active site catalytic residue to form a glycosyl–enzyme adduct. The electron-withdrawing 2-fluoro group stabilizes this adduct allowing for the glycosyl–enzyme adduct to be isolated unlike a native counterpart that is readily cleaved by water molecule.

A reporter group (or a chemical handle) conjugated to the probe facilitates the detection of the retaining glycosidase and the direct identification of a catalytic residue of the enzyme. One example of this type of the probe is 6-azido-2,6-dideoxy-2-fluoro-β-galactosyl fluoride (6Az2FGalF), which has an azide handle at the 6-position of the sugar. This probe was used for functional proteomic analysis of β-galactosidases in bacterial cells.[25] Another example is found in the use of 2-deoxy-2-fluoroxylobioside which has a biotin moiety linked to the sugar through a cleavable disulfide linkage to facilitate affinity isolation process.[26] This probe was employed for identifying a novel retaining β-1,4-glycanase from *Cellulomonas fimi* which hydrolyzes natural and unnatural xylo-type substrates more efficiently than their cello-type counterparts. An important consideration in the design of 2-deoxy-2-fluoroglycoside probes used for the detection of glycosidases is the incorporation of a reporter (or a chemical handle) since modification of hydroxyl groups of the sugar other than those at the anomeric positions may not be tolerated by the enzyme. It should be mentioned that although several examples have been reported to detect glycosidases by using activity-based probes, there is no precedent for the detection of glycosyltransferases by this approach. Thus, chemical probes for glycosyltransferases are highly important.

3.2 Affinity-based probes to detect carbohydrate-binding proteins

Glycan-binding proteins (lectins) are involved in various critical cellular-recognition processes. To fully elucidate the biological implications of glycan–protein recognition events and to develop new carbohydrate-based drugs, it is important

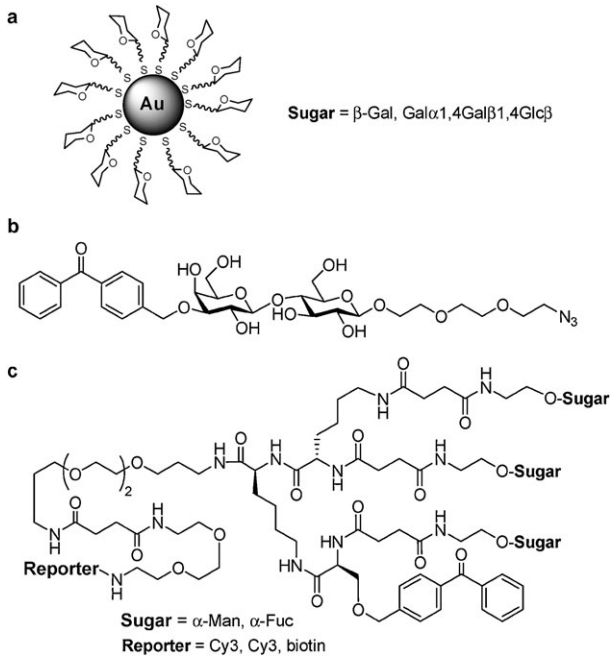

Fig. 6 Affinity-based probes to detect glycan-binding proteins. (a) Carbohydrate-conjugated nanoparticles, (b) trifunctional affinity probes and (c) multivalent trifunctional affinity probes.

to identify and characterize unknown lectins. Since the activity-based approach described above cannot be applied to capture lectins that are devoid of catalytic activity, affinity-based probes have been developed to detect these proteins in complex mixtures.

Carbohydrate-conjugated nanoparticles have been used for studies of carbohydrate–carbohydrate interactions and carbohydrate–protein interactions, detection of pathogens and biolabeling.[27] The carbohydrate–nanoparticle conjugates have been utilized in combination with mass spectrometry (MS) as affinity-based probes for isolation and identification of lectins (Fig. 6a).[28] In this procedure, glycan-conjugated gold nanoparticles are incubated with a mixture of proteins and the protein-bound nanoparticles are separated from the mixture by centrifugation. The isolated protein is identified by determining sequence of peptides obtained from protease digestion by MS analysis.

In addition, trifunctional affinity probes have been developed for the facile detection of lectins. These probes are composed of (1) a carbohydrate ligand, (2) a photoaffinity label for irreversible, covalent labeling of bound lectins, and (3) a reporter or chemoselective ligation motif for visualization and/or isolation. A trifunctional lactose probe was utilized to capture the galectins that bind β-galactoside moieties and are involved in various biological processes and tumor metastasis (Fig. 6b).[29] In this probe, a photoreactive benzophenone moiety is linked to the 3-OH position of the galactose in the lactoside since modification at this site does not affect binding of the sugar to galectins. An azide is linked to the anomeric position of the glucose moiety in the probe through a hydrophilic tether. The galectins that are captured from the protein mixture upon irradiation are subjected to click chemistry using alkyne-modified fluorescent dye for visualization.

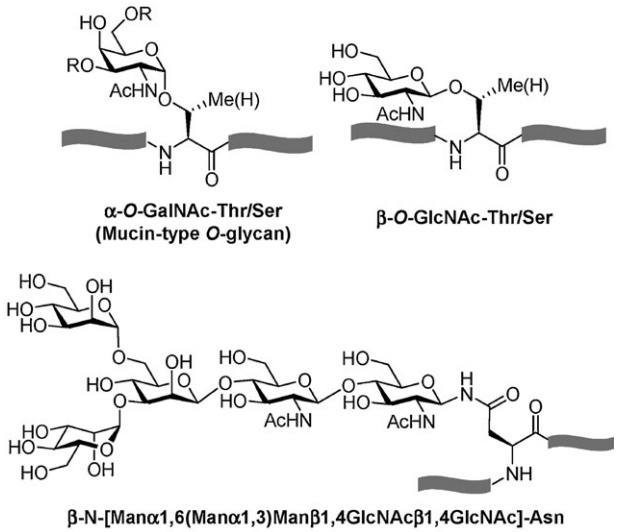

Fig. 7 Structure of common *O*- and *N*-glycans linked to amino acid residues found in glycoproteins.

Multivalent trifunctional affinity probes have been also developed to enhance the binding affinities of glycans to lectins (Fig. 6c).[30] Poor lectin binding properties of the probes caused by modification of the sugar at positions other than the anomeric center can be avoided by attaching the photoreactive benzophenone group to a linker in the probes. The multivalent trifuctional probes can be readily prepared by a safety-catch linker strategy. Labeling studies using these probes demonstrate that specific lectins can be captured with high sensitivity and selectivity. Moreover, the amount of specific lectins present in two different states is rapidly determined by using the probes containing two different fluorescent dyes (*e.g.* Cy3 and Cy5). Overall, the affinity-based probes should be useful in studies of poorly characterized lectins predicted from genomic studies and for profiling the lectins in different cell types or cell states.

4. Synthesis of homogeneous glycoproteins

Glycosylation of proteins in higher organisms is the most common posttranslational modification process because more than 50% of all proteins possess various glycan chains. Glycoproteins carry *N*-glycans linked to an asparagine residue (*N*-linked glycoproteins) and/or *O*-glycans linked to a serine/threonine residue (*O*-linked glycoproteins) of the protein backbone. The most popular form of *O*-glycosylation is mucin-type glycosylation, where *N*-acetylgalactosamine (GalNAc) is attached to a serine/threonine residue of the polypeptide in the α-anomeric configuration (Fig. 7). The C-3 and/or C-6 hydroxyl groups of GalNAc are further glycosylated to produce proteins with more diverse glycans. Another important type of *O*-glycosylation contains a single β-*O*-*N*-acetylglucosamine (GlcNAc) linked to a serine/threonine residue of nuclear and cytoplasmic proteins. *N*-Glycosylation of proteins is more prevalent than their *O*-glycosylation. A core β-$Man_3GlcNAc_2$ oligosaccharide is linked to an asparagine side chain of the polypeptide containing sequence Asn–Xaa–Ser/Thr where Xaa is not proline (Fig. 7). The core pentasaccharide can be

further modified to generate complex, high-mannose and hybrid type glycans. Despite the significance of protein glycosylation, the molecular basis of its function is not well understood owing to difficulties associated with separation of homogeneous glycoproteins from mixtures of glycoforms that contain various glycans in the same polypeptide scaffold. Convergent assembly and fragment assembly are two approaches that have been commonly used to prepare proteins with well-defined glycan structures.

4.1 Glycoprotein synthesis by convergent assembly strategy

The convergent assembly approach involves the direct and site-specific attachment of glycans to the protein backbones or simple glycan-containing proteins by using chemoselective reactions and/or enzymatic reactions. Cysteine is a commonly used site for the modification of the protein. Glycosylation of one cysteine takes place by reacting with thiol-specific carbohydrate reagents, such as maleimide-, selenenylsulfide- and iodoacetamide-conjugated glycans (Fig. 8a).[31,32] The cysteine residue to be glycosylated is often incorporated into the specific site in the protein by using a site-directed mutagenesis technique.

The chemoenzymatic convergent approach, using the endo-β-*N*-acetylglucosaminidase (ENGase), has been developed also to prepare glycoproteins containing homogeneous complex *N*-glycans. Some ENGases, such as Endo-A from *Arthrobacter protophormiae* and Endo-M from *Mucor hiemalis*, have *trans*-glycosylation activity. They can catalyze the transfer of the glycan of a natural *N*-glycopeptide donor to a GlcNAc-peptide acceptor to generate a new glycopeptide. However, ENGase-catalyzed glycosylation has suffered low efficiency of *trans*-glycosylation and should be performed by using only natural *N*-glycopeptides as donor substrates that are difficult to prepare. Recently, an efficient method involving ENGase-catalyzed synthesis of glycopeptides/glycoproteins using oligosaccharide oxazolines as donor substrates was developed.[33,34] Based on the assumption that they are transition state analogs, oligosaccharide oxazolines are reactive glycosyl donors for Endo-A and Endo-M. The glycans from oligosaccharide oxazolines are transferred by ENGases to GlcNAc-containing proteins/peptides in a high yield (Fig. 8b). Notably, ENGases tolerate some modification of the mannose moiety of the Manβ1,4GlcNAc-oxazoline, which is a minimal substrate recognized by these enzymes. This method has been used to obtain homogeneously *N*-glycosylated ribonuclease B (RNase B). The protein was initially treated with Endo-H to remove heterogeneous *N*-glycans, leaving only the innermost GlcNAc unit. The homogeneous GlcNAc-containing RNase B was then treated with oligosaccharide oxazolines in the presence of ENGase (Endo-A) to produce homogeneously *N*-glycosylated RNase B in a high yield (>80%).[33]

Homogeneous glycoproteins are also obtained by using *in vivo* suppressor tRNA technology which allows for site-specific introduction of a selected group into a protein of interest. For example, a ketone-possessing amino acid (*p*-acetyl-L-phenylalanine) is initially incorporated at the amber nonsense codon to produce the ketone-containing mutant protein. Glycoprotein mimetics are then produced by *in vitro* ligation of a ketone moiety in the protein to aminooxy-conjugated carbohydrates (Fig. 8c).[35] In addition, natural glycosylated amino acids (β-GlcNAc-serine and α-GalNAc-threonine) are directly embedded in proteins by incubating

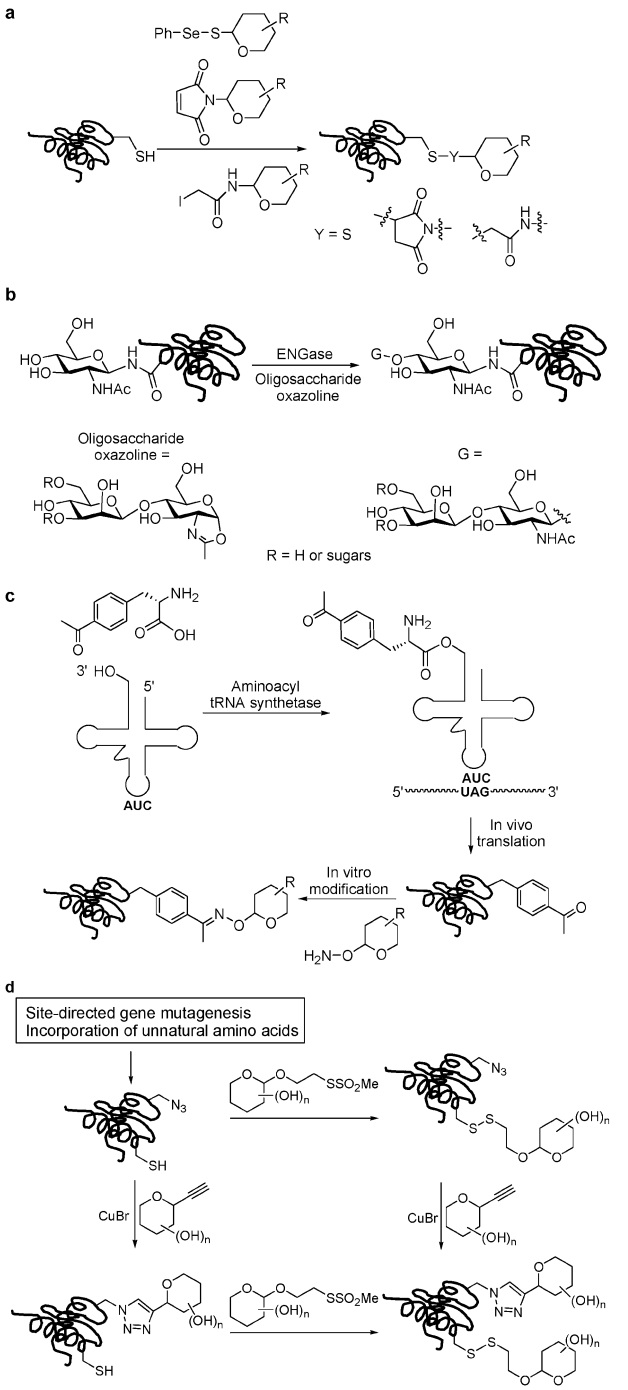

Fig. 8 Glycoprotein synthesis by convergent assembly approach. (a) Attachment of thiol-specific glycans to a cysteine residue of proteins, (b) ENGase-catalyzed *trans*-glycosylation to produce complex *N*-glycoproteins, (c) coupling of aminooxy sugars to the keto-possessing protein obtained from *in vivo* suppressor tRNA technology and (d) dual chemical glycosylation of proteins.

with Ac_3-β-GlcNAc-serine or Ac_3-α-GalNAc-threonine to produce singly glycosylated proteins.[36] The single glycan on the glycoproteins generated in this manner can be further modified by using glycosyltransferases to build more complex carbohydrate appendages.

A majority of naturally occurring glycoproteins contain multiple glycosylations. Thus, a convergent assembly approach to prepare multiply glycosylated proteins is required. Recently, an elegant method to doubly glycosylate a protein scaffold has been developed.[37] To modify the protein with two different glycans, two orthogonal chemical tags (*e.g.* thiol-azide or thiol-alkyne) are required so that the glycans can be sequentially attached to the protein (Fig. 8d). For this purpose, the protein is engineered to contain a single cysteine and a single azidohomoalanine (or homopropargylglycine). All except one cysteine in the native protein are replaced with isosteric serines by site-directed mutagenesis. An azidohomoalanine or homopropargylglycine residue specifically replaces a methionine to obtain the second tag-possessing proteins by using an auxotrophic Met-deficient bacterial strain after the unrequired methionine residues are substituted by isosteric isoleucine by site-directed mutagenesis. The protein containing one cysteine and one azidohomoalanine is first glycosylated by reaction with glycomethanethiosulfonates, cysteine-specific glycan derivatives. In the second step, the azide of azidohomoalanine in the monoglycosylated protein is glycosylated by using Cu(I)-catalyzed glyco-cycloaddition. The reverse order of reactions (*i.e.*, glyco-cycloaddition followed by cysteine-specific glycosylation) also produces the same glycoprotein. Overall, the convergent assembly strategy enables relatively facile preparation of correctly folded glycoproteins. However, this method is difficult to apply for the production of multiply glycosylated proteins with native linkages.

4.2 Glycoprotein synthesis by fragment assembly strategy

The fragment assembly approach involves the chemoselective coupling of two (glyco)peptide fragments to give a glycoprotein with a native amide linkage at the site of ligation. In general, glycopeptides possessing one or more glycan moieties are first prepared by solid-phase peptide synthesis (SPPS). The glycopeptides are then coupled to other (glyco)peptides or recombinant protein fragments to generate glycoproteins by native chemical ligation (NCL) or expressed protein ligation (EPL), respectively.

In NCL approach, (glyco)peptides containing a C-terminal thioester can be synthesized on a solid support by a safety-catch linker strategy (Fig. 9a).[38] Since the thioester linkage on the solid support is not stable under Fmoc deprotection conditions, a safety-catch linker is highly useful to prepare (glyco)peptides containing a C-terminal thioester. The other (glyco)peptides bearing an N-terminal cysteine residue are assembled by conventional SPPS. Upon mixing two fragments, which have no protecting groups on the amino acid side chains but have acetyl groups to protect the hydroxyl groups of the carbohydrates, under ligation conditions, *trans*-thioesterification followed by spontaneous S → N acyl rearrangement occurs chemoselectively to yield glycoproteins with a natural amide linkage at the ligation site. Removal of acetyl protecting groups produces the desired glycoproteins.

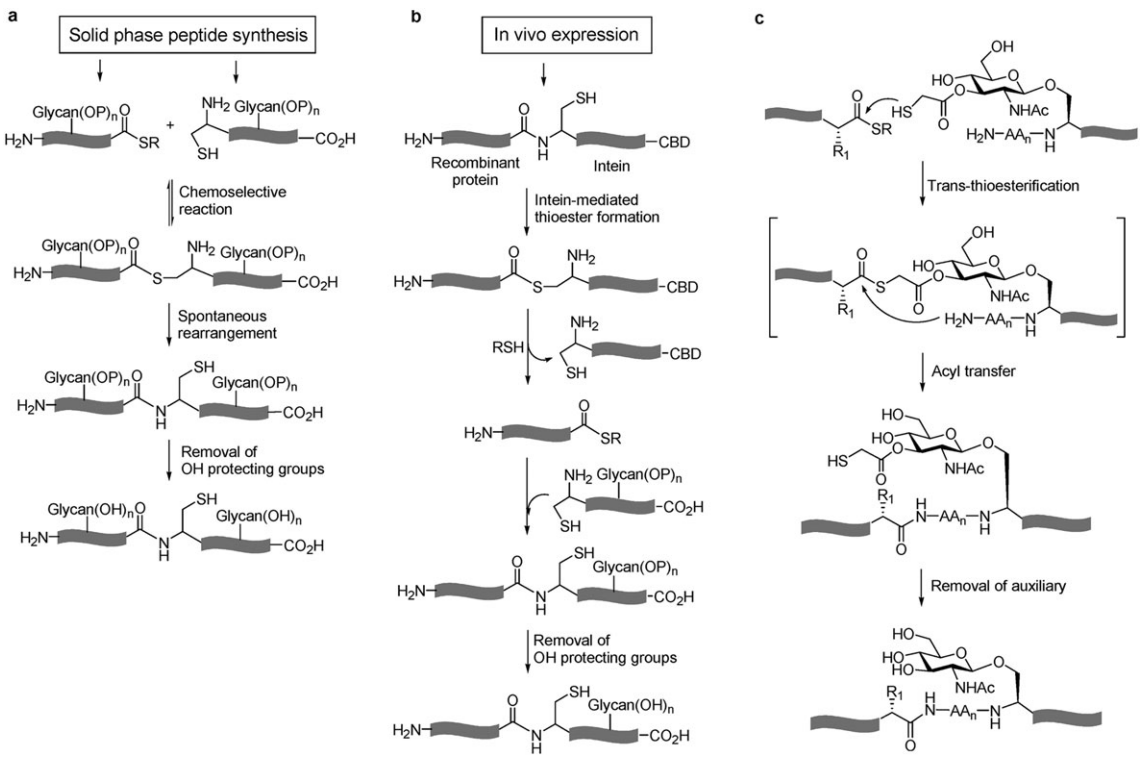

Fig. 9 Glycoprotein synthesis by fragment assembly strategy. (a) Native chemical ligation, (b) expressed protein ligation and (c) sugar assisted ligation (P: hydroxyl protecting groups).

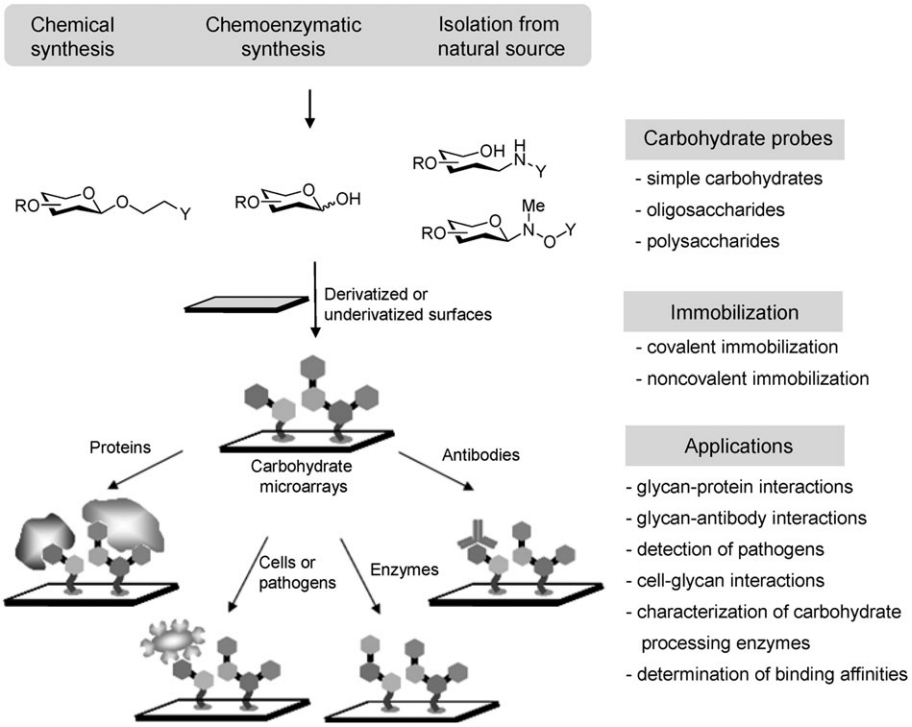

Fig. 10 Fabrication of carbohydrate microarrays and their applications for biological and biomedical research.

Several glycoproteins and glycoprotein fragments, such as diptericin (an antibacterial glycoprotein with 82 amino acid residues), lymphotactin (a 93-residue chemokine containing eight *O*-glycosylation sites), prostate-specific antigen (24–47) and RNA B (40–68) fragments containing complex *N*-glycans, have been prepared by using this approach.[39]

Although the NCL approach is suitable for the preparation of relatively small glycoproteins, it is inappropriate for the generation of large glycoproteins owing to the fact that only peptides with less than 60 amino acid residues can be obtained from SPPS. Large glycoproteins can be prepared by coupling synthetic glycopeptides to peptide fragments produced recombinantly in *E. coli* in a process known as EPL (Fig. 9b). The advantage of this method is that it is possible to obtain much larger coupling fragments in a cost-effective way than with SPPS. The procedure begins with preparation of a sugar-protected glycopeptide containing an *N*-terminal cysteine residue by using conventional SPPS. The other large peptide is expressed as an intein fusion in *E. coli*. The peptide–intein fusion undergoes the intein-mediated *trans*-thioesterification. Upon treatment of the fusion proteins with thiol-reactive reagents such as dithiothreitol, a protein fragment containing a *C*-terminal thioester is released and then ligated to the synthetic glycopeptide in the presence of 2-mercaptoethanesulfonic acid to give the sugar-protected glycoproteins. Removal of sugar-hydroxyl protecting groups produces the desired glycoproteins. In EPL, chitin-binding domain (CBD) fused to intein is used for the facile separation of protein obtained after expression in *E. coli* and cleavage of intein–CBD fusion. By using the EPL approach, three glycoforms of GlyCAM-1, a mucin-like glycoprotein with 132 amino acid residues that acts as a ligand for the leukocyte adhesion molecule L-selectin, have been successfully obtained.[40]

Both the NPL and EPL methods require the presence of an *N*-terminal cysteine residue for capturing the *C*-terminal thioester peptide or the protein fragment. Because of the scarcity of cysteine residues in natural proteins, cysteine-free chemical ligation methods have been developed. One interesting method uncovered involves sugar assisted ligation.[41] In this approach, *trans*-thioesterification takes place between the *C*-terminal glycopeptide containing GlcNAc that is derivatized with a mercaptoacetate auxiliary at the 3-OH and *N*-terminal (glyco)peptide with a thioester group (Fig. 9c). Subsequent S → N acyl transfer occurs, like in NCL, to produce a glycopeptide. Removal of the mercaptoacetate group generates the desired glycopeptide. Importantly, the sugar assisted ligation technique tolerates a number of possible ligation junctions and the auxiliary can be located up to five amino acids away from the ligation site without deleterious effects on ligation efficiency. Even though full-length synthesis of polypeptides containing complex glycans has not been accomplished thus far, the preparation of complex glycan-linked glycoproteins may be possible by using this method.

Overall, the fragment assembly strategy enables production of multiply glycosylated proteins with native linkages. However, the glycoproteins prepared after fragment ligation should be properly folded in order to reveal their full bioactivities. The sustained effort made to produce homogeneous glycoproteins by using either convergent assembly strategy or fragment assembly strategy will ultimately provide a greater understanding of the molecular basis of the protein glycosylation and lead to the development of glycoprotein-based therapeutic agents.

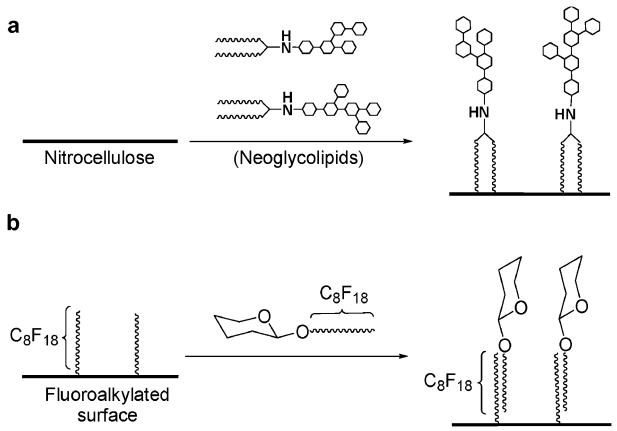

Fig. 11 Noncovalent, site-specific immobilization of (a) glycolipids and (b) fluoroalkylated sugars on nitrocellulose membranes and fluoroalkylated surfaces, respectively.

5. Carbohydrate microarrays

Glycan–protein interactions are involved in a variety of physiological and pathological processes. Conventional approaches, including the hemagglutination inhibition assay, enzyme-linked lectin assay, surface plasmon resonance and isothermal titration calorimetry, have been widely used to evaluate glycan–protein recognition events over past decades. Although successfully applied for studying these biomolecular interactions, the conventional techniques are labor intensive and often require large amounts of carbohydrate samples. As a high-throughput analytic tool for glycan–protein interactions, carbohydrate microarrays, which are composed of diverse glycans densely attached to the solid surface in an orderly arrangement, was for the first time reported in 2002.[42–45] The notable advantage of the carbohydrate microarray-based technology is that in contrast to most conventional methods it allows for simultaneous analysis of a number of glycan–protein interactions using small amounts of carbohydrate samples.

Carbohydrate microarrays are prepared by immobilizing modified or unmodified carbohydrates on the proper solid surfaces (Fig. 10).[46,47] The glass slide is the most extensively used surface material but gold and nitrocellulose membrane are also utilized for this purpose. The solid surfaces, in particular, glass slides, are properly derivatized for the efficient and selective attachment of carbohydrates. The required carbohydrates (simple carbohydrates, oligosaccharides and polysaccharides) are obtained from chemical synthesis, chemoenzymatic synthesis, and a natural source such as glycoprotein and glycolipid glycans. Advances in chemical glycosylation (*e.g.*, automated oligosaccharide synthesis, one-pot oligosaccharide synthesis and combinatorial carbohydrate synthesis) make the use of chemical synthetic methods more popular for the preparation of diverse carbohydrate probes.[3]

To date, various immobilization strategies for carbohydrates on solid surfaces have been developed. For example, lipid and fluorescent tag-conjugated oligosaccharides are noncovalently, site-specifically adsorbed to nitrocellulose/polyvinylidene difluoride membranes and fluoroalkylsilane-coated glass slides, respectively (Fig. 11).[44,48] Unmodified polysaccharides are immobilized on the solid surface through nonspecific adsorption.[43] However, a great effort has been given to the construction of carbohydrate microarrays employing glycans linked to specific functional groups that selectively react with functional groups on the surface to form covalent attachments. For instance, maleimide, hydrazide, aminooxy, thiol, amine and diene-conjugated glycans have been linked to thiol, epoxide, aldehyde, maleimide, *N*-hydroxysuccinimide (NHS) and benzoquinone-coated surfaces (Fig. 12).[46,49]

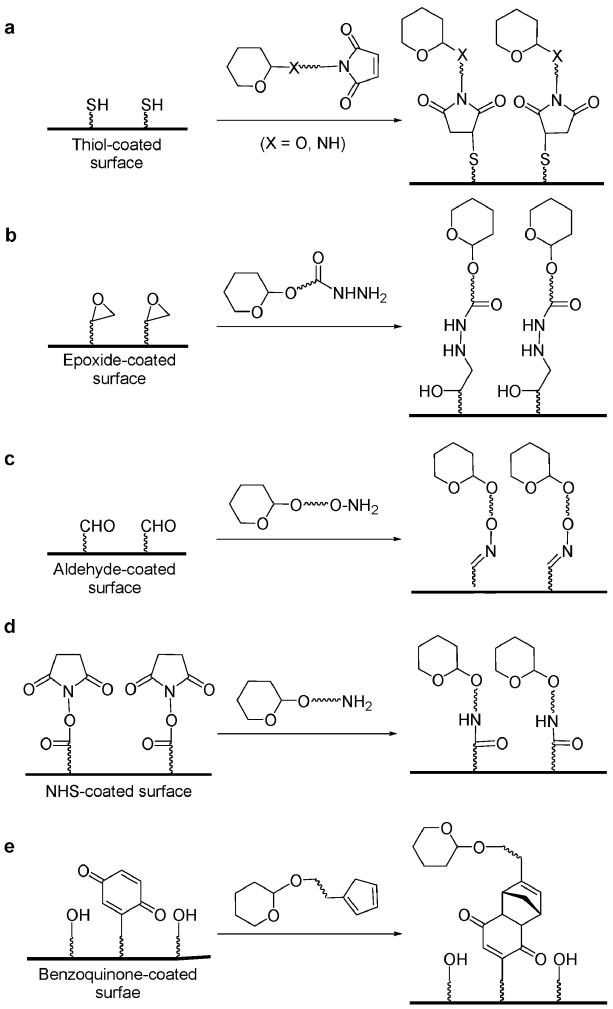

Fig. 12 Covalent, site-specific immobilization of (a) maleimide, (b) hydrazide, (c) aminooxy, (d) amine and (e) cyclopentadiene-conjugated sugars on properly derivatized surfaces.

In many cases, functionalized glycans are prepared by multistep sequences. In order to circumvent the labor-intensive and time-consuming nature of these procedures, one-step methods for the modification of carbohydrates with functional groups required for covalent immobilization have been exploited. In one case, free carbohydrates are reacted with 2,6-diaminopyridine in the presence of sodium cyanoborohydride or an *N*-methylaminooxy-containing bifunctional linker to produce acyclic or cyclic adducts, respectively (Fig. 13a).[50,51] The modified glycans appended by the amine group are then immobilized on NHS-activated glass slides.

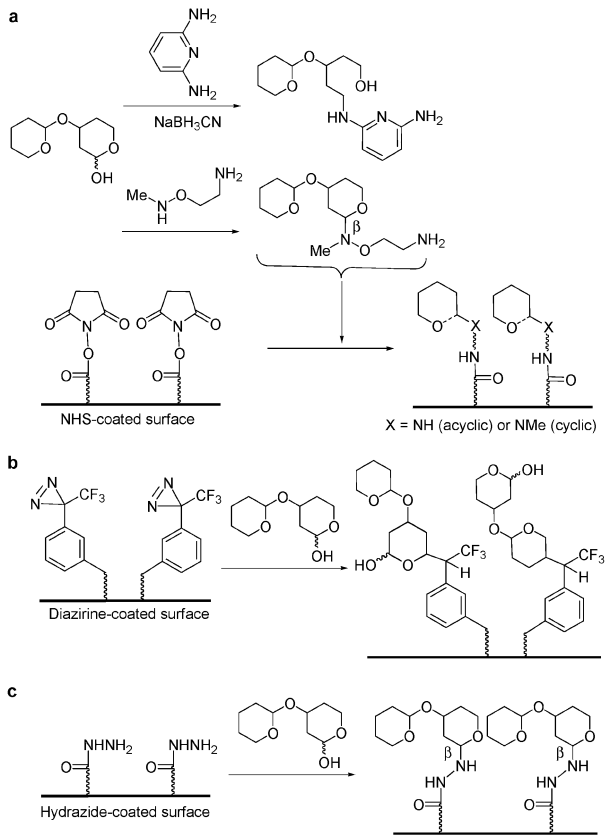

Fig. 13 Immobilization of (a) carbohydrates obtained by one-step reaction on NHS-coated surface, (b) free glycans on photolabile group-coated surface, and (c) free glycans on hydrazide-coated surface.

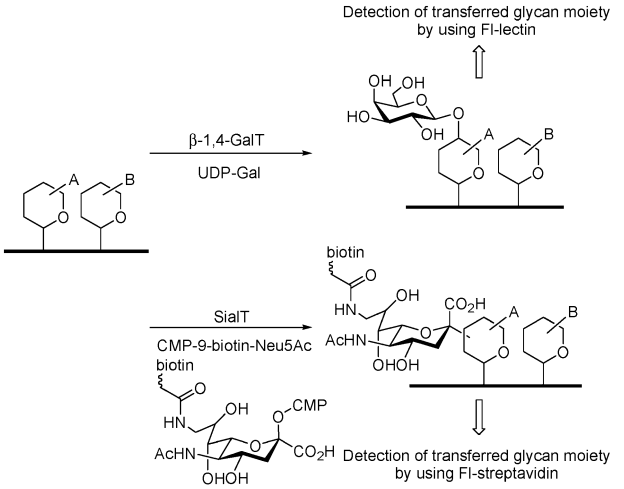

Fig. 14 Characterization of glycosyltransferase activities by using carbohydrate microarrays (A, B; different sugars immobilized on the solid surface, GalT; galactosyltransferase, SialT; sialyltransferase).

However, modification of variously sized carbohydrates by either one step or multistep routes is still a hurdle to the fabrication of microarrays containing diverse carbohydrates. In order to circumvent the need to use modified glycans, the immobilization methods that are applicable to unmodified carbohydrates have been developed. One approach involves the use of photolabile group-coated glass slides to covalently immobilize unmodified carbohydrates.[52] For example, photolabile aryltrifluoromethyl-diazirine groups on the surface are transformed to reactive carbenes upon irradiation. The generated carbene species react with free glycans in their vicinity to form covalent bonds (Fig. 13b). A drawback of this method is the nonspecific attachment of carbohydrates to the surface owing to the nonspecific nature of the carbene reactions. An alternative immobilization method, in which variously sized unmodified carbohydrates (simple carbohydrates, oligosaccharides and polysaccharides) are site-specifically and covalently immobilized on hydrazide-coated glass slides, has been developed (Fig. 13c).[53] In this process, the reducing sugars of glycans are attached to the hydrazide group on the surface in the cyclic form with β-configuration. Importantly, immobilization methods that rely on free carbohydrates can be used even by biologists who lack organic synthesis experience.

Since its advent in 2002, the glycan microarray technology has been applied at an increasing rate in the areas of biological and biomedical research (Fig. 10). The most extensive use of this technology has been in the rapid analysis of the specificities of various carbohydrate-binding proteins including lectins, antibodies, cytokines, chemokines and growth factors. The binding properties of proteins are examined by treating the carbohydrate microarrays containing diverse glycans with fluorophore-labeled proteins or unlabeled proteins and subsequent incubation with labeled antibodies. The results obtained by using carbohydrate microarrays provide information about the detailed binding properties of various plant and animal lectins.[46]

Glycosaminoglycans (GAGs), such as chondroitin sulfate, heparin, heparan sulfate, keratan sulfate, dermatan sulfate and hyaluronan, play roles in various physiological processes (*e.g.*, homeostasis, cancer metastasis, cell growth, cell migration and development) through interactions with proteins. Because of the biological significance of GAGs, carbohydrate microarrays immobilized by synthetic chondroitin sulfate and heparin oligosaccharides have been fabricated and applied to the rapid determination of binding specificities of GAG-binding proteins, such as growth factors, cytokines and chemokines, and for profiling the sulfation specificities of GAG interactions with proteins.[54,55] Studies of the latter process have demonstrated that specific sulfation motifs of oligosaccharides are critical for their binding to growth factors.

Carbohydrate microarrays are useful in the investigation of the binding properties of anti-glycan antibodies. Profiling of glycan–antibody interactions by using carbohydrate microarrays has shown that many anti-glycan antibodies, which are considered to be specific for their designated glycan antigens, cross-react with other glycans.[56] Therefore, the results obtained from the use of anti-glycan antibodies should be carefully analyzed and interpreted. The microarray technology has been also employed to study mammalian cell–glycan interactions through binding of cell surface proteins to glycans.[57]

Carbohydrate microarrays can be applied in biomedical investigations such as for the detection of disease-related anti-glycan antibodies and pathogens.[46] Pathogens express specific immunogenic glycans on their surfaces and

pathogen-infected humans elicit antibodies that bind to the pathogenic glycans. Detection of these antibodies using glycan microarrays is a useful method for diagnosis of diseases. In addition, the microarray technology can be employed to directly detect intact bacteria and viruses, and bacterial toxins.[46] Furthermore, carbohydrate microarrays have been applied to address biological issues such as the identification of HIV vaccine candidate antigens.

The microarray technology can also be used for the rapid analysis of acceptor specificities of glycosyltransferases and for measurement of the relative rates of glycosylation of substrates by these enzymes. In this process, glycan microarrays are treated with glycosyltransferases in the presence of either natural glycosyl donor (*e.g.* UDP-Gal) or biotin-labeled modified glycosyl donor (*e.g.* CMP-9-biotin-Neu5Ac) followed by detection of the transferred sugar moiety by using specific lectin or fluorophore-labeled streptavidin, respectively (Fig. 14).[58,59] Another interesting application of carbohydrate microarrays involves the quantitative analysis of protein-binding affinities. Initially, IC_{50} values of soluble inhibitors for protein binding to carbohydrates immobilized on the surface were determined by using this technology.[58,60] More recently, various K_d values (dissociation constants) between proteins and surface-immobilized carbohydrates were measured with a single experiment by using carbohydrate microarrays.[58,60] Dissociation constants for surface-bound sugars with proteins obtained from microarray experiments were found to be similar to those obtained in SPR experiments.

As demonstrated here, carbohydrate microarrays are suitable for studies of a wide variety of glycan-mediated biological processes. It would be ideal if microarrays containing the entire glycome of an organism on a single slide could be constructed. For this purpose, further development of efficient synthetic methods to prepare diverse carbohydrate probes is needed. Continued advances in this area will make the carbohydrate microarray technology a more general and practicable platform for functional studies of glycans.

6. Conclusions

For many years, researchers studying the functions of glycans have suffered from the lack of proper tools unlike those who study the functional roles of proteins or nucleic acids. As a consequence, functional studies of glycans present in proteins, cells and organisms at the molecular level have lagged behind in comparison with functional studies of proteins. Over the past few years, many chemical approaches have been rapidly exploited to aid in the understanding of biological roles played by glycans. However, these advances have still not led to an advanced understanding of the functions of glycans. The challenges that remain in the area of glycans relate not only to the development of efficient, powerful chemical methods but also to applications of the tools in addressing biological questions. Continuing efforts should provide a deeper insight into the molecular basis of the function of glycans in biological systems. Moreover, these advances should also lead to the development of new carbohydrate-based therapeutic agents.

Acknowledgements

We apologize to the many authors whose excellent papers could not be cited in this review for reference limitations. This work was supported by grants from the NRL (M10500000027-O6J0000-02710) and Protein Chip Technology programs of MOST/KOSEF. S.P. thanks the BK21 program (KRF).

References

1. A. Varki and M. J. Chrispeels, *Glycobiology*, 1993, **3**, 97–130.
2. J. A. Prescher and C. R. Bertozzi, *Cell*, 2006, **126**, 851–854.
3. O. J. Plante, E. R. Palmacci and P. H. Seeberger, *Science*, 2001, **291**, 1523–1527.
4. J. Lowe and J. Marth, *Annu. Rev. Biochem.*, 2003, **72**, 643–691.
5. L. K. Mahal, K. J. Yarema and C. R. Bertozzi, *Science*, 1997, **276**, 1125–1128.
6. A. Varki, *Nature*, 2007, **446**, 1023–1029.
7. G. A. Lemieux, K. J. Yarema, C. L. Jacobs and C. R. Bertozzi, *J. Am. Chem. Soc.*, 1999, **121**, 4278–4279.
8. E. Saxon and C. R. Bertozzi, *Science*, 2000, **287**, 2007–2010.
9. J. A. Prescher, D. H. Dube and C. R. Bertozzi, *Nature*, 2004, **430**, 873–877.
10. D. J. Vocadlo, H. C. Hang, E.-J. Kim, J. A. Hanover and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 9116–9121.
11. A. Nandi, R. Sprung, D. K. Barma, Y. Zhao, S. C. Kim, J. R. Falck and Y. Zhao, *Anal. Chem.*, 2006, **78**, 452–458.
12. H. C. Tai, N. Khidekel, S. B. Ficarro, E. C. Peters and L. C. Hsieh-Wilson, *J. Am. Chem. Soc.*, 2004, **126**, 10500–10501.
13. N. Khidekel, S. Arndt, N. Lamarre-Vincent, A. Lippert, K. G. Poulin-Kerstien, B. Ramakrishnan, P. K. Qasba and L. C. Hsieh-Wilson, *J. Am. Chem. Soc.*, 2003, **125**, 16162–16163.
14. N. Khidekel, S. B. Ficarro, P. M. Clark, M. C. Bryan, D. L. Swaney, J. E. Rexach, Y. E. Sun, J. J. Coon, E. C. Peters and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2007, **3**, 339–348.
15. H. C. Hang, C. Yu, D. L. Kato and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 14846–14851.
16. D. H. Dube, J. A. Prescher, C. N. Quang and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 4819–4824.
17. T.-L. Hsu, S. R. Hanson, K. Kishikawa, S.-K. Wang, M. Sawa and C.-H. Wong, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 2614–2619.
18. J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 16793–16797.
19. M. J. Evans and B. F. Cravatt, *Chem. Rev.*, 2006, **106**, 3279–3301.
20. S. Halazy, V. Berges, A. Ehrhard and C. Danzin, *Bioorg. Chem.*, 1990, **18**, 330–344.
21. M. Ichikawa and Y. Ichikawa, *Bioorg. Med. Chem. Lett.*, 2001, **11**, 1769–1773.
22. C.-S. Tsai, Y.-K. Li and L.-C. Lo, *Org. Lett.*, 2002, **4**, 3607–3610.
23. C. P. Lu, C.-T. Ren, Y.-N. Lai, S.-H. Wu, W.-M. Wang, J.-Y. Chen and L.-C. Lo, *Angew. Chem., Int. Ed.*, 2005, **44**, 6888–6892.
24. M. Kurogochi, S.-I. Nishimura and Y. C. Lee, *J. Biol. Chem.*, 2004, **279**, 44704–44712.
25. D. J. Vocadlo and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2004, **43**, 5338–5342.
26. O. Hekmat, Y.-W. Kim, S. J. Williams, S. He and S. G. Withers, *J. Biol. Chem.*, 2005, **280**, 35126–35135.
27. J. M. de la Fuente and S. Penadés, *Biochim. Biophys. Acta*, 2006, **1760**, 636–651.
28. Y.-J. Chen, S.-H. Chen, Y.-Y. Chien, Y.-W. Chang, H.-K. Liao, C.-Y. Chang, M.-D. Jan, K.-T. Wang and C.-C. Lin, *ChemBioChem*, 2005, **6**, 1169–1173.
29. L. Ballell, K. J. Alink, M. Slijper, C. Versluis, R. M. J. Liskamp and R. J. Pierters, *ChemBioChem*, 2005, **6**, 291–295.
30. M.-R. Lee, D. W. Jung, D. Williams and I. Shin, *Org. Lett.*, 2005, **7**, 5477–5480.
31. I. Shin, H.-J. Jung and M.-R. Lee, *Tetrahedron Lett.*, 2001, **42**, 1325–1328.

32. B. G. Davis, *Chem. Rev*., 2002, **102**, 579–601.
33. B. Li, H. Song, S. Hauser and L.-X. Wang, *Org. Lett*., 2006, **8**, 3081–3084.
34. T. W. D. F. Rising, T. D. W. Claridge, J. W. B. Moir and A. J. Faibanks, *ChemBioChem*, 2006, **7**, 1177–1180.
35. H. Liu, L. Wang, A. Brock, C.-H. Wong and P. G. Schultz, *J. Am. Chem. Soc*., 2003, **125**, 1702–1703.
36. Z. Zhang, J. Gildersleeve, Y.-Y. Yang, R. Xu, J. A. Loo, S. Uryu, C.-H. Wong and P. G. Schultz, *Science*, 2004, **303**, 371–373.
37. S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony and B. G. Davis, *Nature*, 2007, **446**, 1105–1109.
38. Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman and C. R. Bertozzi, *J. Am. Chem. Soc*., 1999, **121**, 11684–11689.
39. S. Mezzato, M. Schaffrath and C. Unverzagt, *Angew. Chem., Int. Ed*., 2005, **44**, 1650–1654.
40. D. Macmillan and C. R. Bertozzi, *Angew. Chem., Int. Ed*., 2004, **43**, 1355–1359.
41. S. Ficht, R. J. Payne, A. Brik and C.-H. Wong, *Angew. Chem., Int. Ed*., 2007, **46**, 5975–5979.
42. S. Park and I. Shin, *Angew. Chem., Int. Ed*., 2002, **41**, 3180–3182.
43. D. Wang, S. Liu, B. J. Trummer, C. Deng and A. Wang, *Nat. Biotechnol*., 2002, **20**, 275–281.
44. S. Fukui, T. Feizi, C. Galustian, A. M. Lawson and W. Chai, *Nat. Biotechnol*., 2002, **20**, 1011–1017.
45. B. T. Houseman and M. Mrksich, *Chem. Biol*., 2002, **9**, 443–454.
46. I. Shin, J. Tae and S. Park, *Curr. Chem. Biol*., 2007, **1**, 187–199.
47. J. C. Paulson, O. Blixt and B. E. Collins, *Nat. Chem. Biol*., 2006, **2**, 238–248.
48. K. S. Ko, F. A. Jaipuri and N. L. Pohl, *J. Am. Chem. Soc*., 2005, **127**, 13162–13163.
49. S. Park, M.-R. Lee, S. J. Pyo and I. Shin, *J. Am. Chem. Soc*., 2004, **126**, 4812–4819.
50. B. Xia, Z. S. Kawar, T. Ju, R. A. Alvarzev, G. P. Sachdev and R. D. Cummings, *Nat. Methods*, 2005, **2**, 845–850.
51. O. Bohorov, H. Andersson-Sand, J. Hoffmann and O. Blixt, *Glycobiology*, 2006, **16**, 21C–27C.
52. S. Angeloni, J. L. Ridet, N. Kusy, H. Gao, F. Crevoisier, S. Guinchard, S. Kochhar, H. Sigrist and N. Sprenger, *Glycobiology*, 2005, **15**, 31–41.
53. M.-R. Lee and I. Shin, *Org. Lett*., 2005, **7**, 4269–4272.
54. J. L. de Paz, E. A. Moseman, C. Noti, L. Polito, U. H. von Andrian and P. H. Seeberger, *ACS Chem. Biol*., 2007, **2**, 735–744.
55. C. I. Gama, S. E. Tully, N. Sotogaku, P. M. Clark, M. Rawat, N. Vaidehi, W. A. Goddard III, A. Nishi and L. C. Hsieh-Wilson, *Nat. Chem. Biol*., 2006, **2**, 467–473.
56. J. C. Manimala, T. A. Roach, Z. Li and J. C. Gildersleeve, *Glycobiology*, 2007, **17**, 17C–23C.
57. L. Nimrichter, A. Gargir, M. Gortler, R. T. Altstock, A. Shtevi, O. Weisshaus, E. Fire, N. Dotan and R. L. Schnaar, *Glycobiology*, 2004, **14**, 197–203.
58. S. Park and I. Shin, *Org. Lett*., 2007, **9**, 1675–1678.
59. O. Blixt, K. Allin, O. Bohorov, X. Liu, H. Andersson-Sand, J. Hoffmann and N. Razi, *Glycoconjugate J*., 2008, **25**, 59–68.
60. P.-H. Liang, S.-K. Wang and C.-H. Wong, *J. Am. Chem. Soc*., 2007, **129**, 11177–11184.