Carbohydrates as the Next Frontier in Pharmaceutical Research

DOI: 10.1002/chem.200500025 *Chem. Eur. J.* **2005**, *11*, 3194 – 3206

Carbohydrates as the Next Frontier in Pharmaceutical Research

Daniel B. Werz and Peter H. Seeberger*[a]

Abstract: Synthetic carbohydrates and glycoconjugates are used to study their roles in biological important processes such as inflammation, cell–cell recognition, immunological response, metastasis, and fertilization. The development of an automated oligosaccharide synthesizer greatly accelerates the assembly of complex, naturally occurring carbohydrates as well as chemically modified oligosaccharide structures and promises to have major impact on the field of glycobiology. Tools such as microarrays, surface plasmon resonance spectroscopy, and fluorescent carbohydrate conjugates to map interactions of carbohydrates in biological systems are presented. Case studies of the successful application of carbohydrates as active agents are discussed, for example, fully synthetic oligosaccharide vaccines to combat tropical diseases (e.g., malaria), bacterial infections (e.g., tuberculosis), viral infections such as HIV, and cancer. Aminoglycosides serve as examples of drugs acting through carbohydrate–nucleic-acid interactions, while heparin works by carbohydrate–protein interactions. A general, modular strategy for the complete stereoselective synthesis of defined heparin oligosaccharides is presented. A carbohydrate-functionalized fluorescent polymer has been shown to detect miniscule amounts of bacteria faster than commonly used methods.

Keywords: automated synthesis · carbohydrates · glycobiology · glycoconjugates · heparin · microarrays

[a] Dr. D. B. Werz, Prof. Dr. P. H. Seeberger
Laboratory for Organic Chemistry
Swiss Federal Institute of Technology Zürich, ETH-Hönggerberg
HCI F315, Wolfgang-Pauli-Strasse 10, 8093 Zürich (Switzerland)
Fax: (+41) 163-31-235
E-mail: seeberger@org.chem.ethz.ch

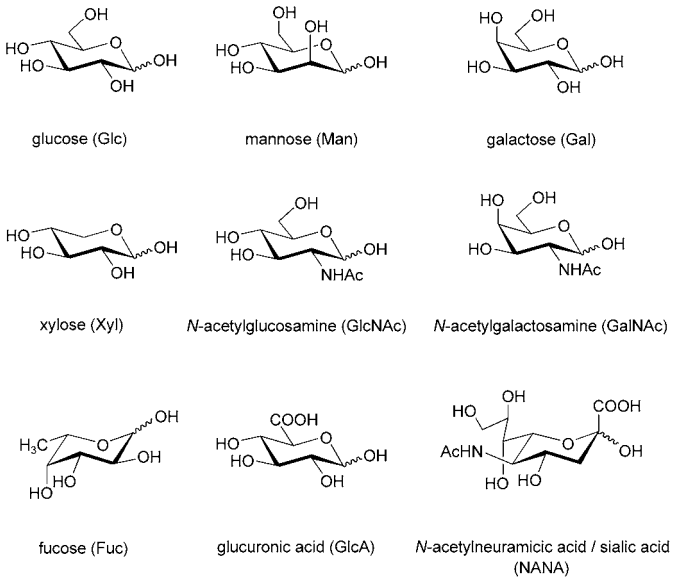
Introduction

Three major classes of polymers are responsible for the storage of information and signal transduction processes in biological systems: these are nucleic acids, proteins, and polysaccharides. DNA, the genetic material transferring information from generation to generation, functions as the blueprint of life. RNA serves as a transient repository of genetic information on the way from DNA to proteins, but also plays pivotal roles in cell division, gene expression, and catalysis. The protein-synthesis machinery, called the ribosome, consists of RNA.[1] Proteins, the second major class of biopolymers, are encoded by nucleic acids and constitute the catalytic machinery carrying out most of the reactions in the cell. Proteins are also important as skeletal material of numerous organisms to provide strength as well as flexibility. Glycosyltransferases, a special class of enzymes, are responsible for the synthesis of carbohydrates, the third class of biopolymers.

While nucleic acids and proteins are linear assemblies, carbohydrates are the most complex and diverse class of biopolymers commonly found in nature as glycoconjugates. A wide array of available monosaccharide building blocks as well as the possibility of different stereochemical linkages between each pair of carbohydrates results in tremendous complexity. Additionally, the chain length of the oligosaccharides can also vary widely from monosaccharides up to branched oligosaccharides with more than 30 building blocks, or in the case of polysaccharides to several thousand building blocks. The most prominent examples for the latter type are cellulose, the major constituent of plant tissues, and chitine, which forms the shells of insects and crabs. The nine common monosaccharides found in mammalian cells (shown here) can be combined in a dazzling variety of ways to form structures more diverse than those accessible with the twenty naturally occurring amino acids or the four nucleotides.

Moreover, oligosaccharides are present in the form of glycoconjugates (glycoproteins and glycolipids) in all cell walls mediating a variety of events such as inflammation, cell–cell recognition, immunological response, metastasis, and fertil-

 DOI: 10.1002/chem.200500025



ization.[2] The carbohydrate coat surrounding a cell called the glycocalix is specific for a particular species, cell type, and developmental status. Alterations in cell surface oligosaccharides have been found in association with many pathological conditions such as cancer and tuberculosis.

Usually, the desired glycoconjugates exist in heterogeneous mixtures that are difficult to isolate in pure form, and when at all possible, only small amounts are obtained. For the other two major classes of biopolymers many tools are available to elucidate their structure, their function, and their structure–function relationships. Detailed insights have been gained into protein–protein interactions, protein–nucleic-acid interactions and nucleic-acid–nucleic-acid interactions. This research has been of fundamental importance for the development of new therapeutics that aim to modify, enhance, or disrupt these interactions. In contrast, carbohydrates, although studied for more than 100 years, have attracted less interest in the field of drug discovery. Forty years ago, biochemical research concerning carbohydrates was focused on their role in energy storage and supply in biological systems. Biosynthesis and biodegradation pathways were discovered; however, the function of carbohydrates in biologically important recognition processes and their potential use as drugs became evident much later. Thus, all aspects of glycobiology, now often termed glycomics, are still less well understood than its two counterparts genomics and proteomics, dealing with nucleic acids and proteins.

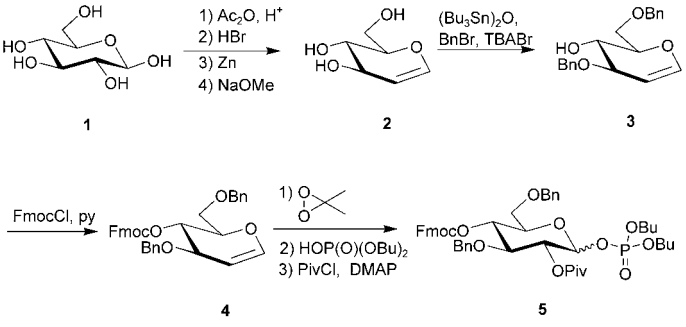
Two major technological breakthroughs have catapulted peptide- and oligonucleotide-based research forward: the sequencing of nucleic acids and proteins has been automated and allows for the composition of an unknown sample to be determined quickly and reliably. Secondly, the synthesis of defined oligonucleotides[3] and peptides[4] also has been automated and even allows nonspecialists in this field to obtain rapidly larger-scale quantities of these important classes of biopolymers. The rational design of specific modifications is possible and has been used successfully to create important research tool for biomedicine, biotechnology, and pharmaceutical areas.

In contrast, oligosaccharide sequencing and structure determination remains a difficult task, even though major efforts have been directed towards the improvement of modern analytical methods such as HPLC, two-dimensional NMR techniques, and special mass spectroscopic methods, such as ESI and MALDI-TOF.[5] Until recently, access to pure oligosaccharides remained technically difficult and extremely time-consuming. Multiple chemical[6] and enzymatic methods[7] are known, and finally an automated method has been developed.[8]

Automated Carbohydrate Synthesis

Analogous to the highly efficient synthesis of peptides and oligonucleotides, solid-phase synthesis has been used for the automated assembly of oligosaccharides.[8,9] Two advantages of the solid-phase approach are noteworthy: 1) the use of excess reagent drives reactions to completion and 2) purification after each reaction step is not required, but rather washing procedures remove excess reagents.[8,9]

Our laboratory decided to utilize an acceptor-bound approach for the carbohydrate assembly, whereby the anomeric position of the first carbohydrate is attached at its reducing end to the solid support.[8,9] Therefore, glycosyl phosphates[10] and glycosyl trichloroacetimidates[11] proved to be ideal glycosylating agents that are relatively stable and can be stored for many months in the refrigerator. Glycosyl phosphates are readily synthesized by a one-pot procedure starting from differentially protected glycals (Scheme 1).



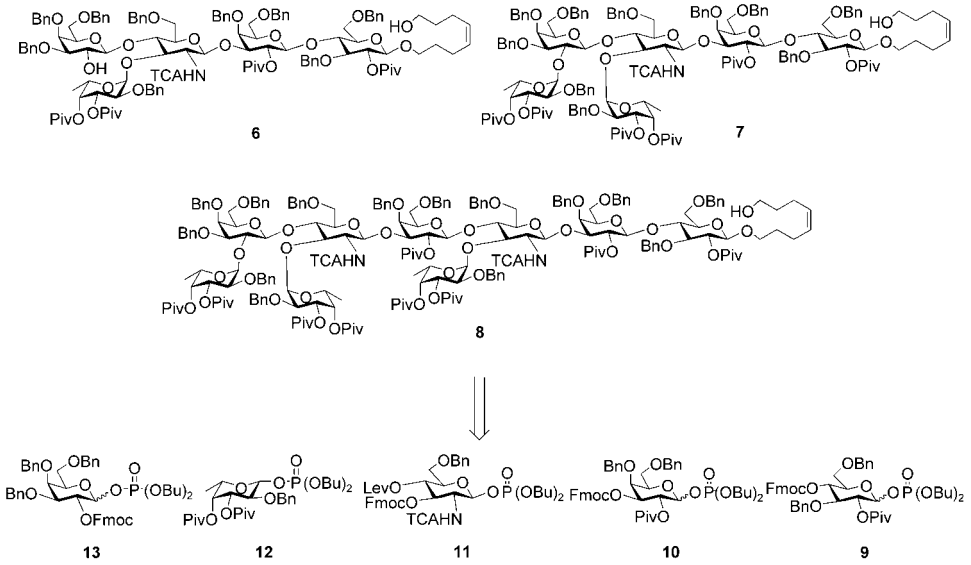
Scheme 1. Representative synthesis of glycosyl phosphate **5** through the epoxidation of glycal **4**.

Firstly, a Zn-mediated reduction eliminates two hydroxyl groups of the glucose **1** to afford the glycal **2**. Protection with permanent (Bn) and temporary (Fmoc) protecting groups yield **3** and **4**, respectively. Epoxidation with dimethyl dioxirane (DMDO) is followed by opening of the 1,2-anhydrosugar with dibutylphosphate. Protection of the ensuing

C2 hydroxyl group (e.g., by using pivaloyl chloride) produced the desired glycosyl phosphate **5** in excellent yield.[10]

Glycosylation reactions in the presence of trimethylsilyl triflate result in good yields. The reaction times usually range between 10 and 30 minutes. Selectivity at the anomeric center is achieved by using appropriate participating, for example, pivaloyl (Piv), benzoyl (Bz), 9-fluorenylmethoxycarbonyl (Fmoc) or nonparticipating groups, for example, benzyl (Bn), at the C2 hydroxyl group. Easily and selectively removable temporary protecting groups such as Fmoc, which is cleaved by weak bases such as piperidine or triethylamine, have been shown to be important for successful oligosaccharide syntheses.[12] Orthogonal protecting groups are utilized in concert to access branched oligosaccharides.[12,13] Following selection of a flexible and useful protecting group strategy, the next strategic consideration involves the choice of an appropriate resin and the right linker connecting the first sugar at its reducing end with the solid support. The linker has to be compatible with a wide range of reaction conditions applied during oligosaccharide assembly. However, after the synthesis is completed, rapid and efficient cleavage is necessary. Two linkers that are readily connected to Merrifield's resin have shown to fulfill these requirements: 1) an alkene-containing linker,[14] which is released from the solid support by olefin cross-metathesis by using Grubbs' catalyst and ethylene, and 2) an ester-containing linker (**14**), which is cleaved by strong bases such as methoxide.[12] The latter linker can only be used when the deprotecting conditions during oligosaccharide assembly avoid strong basic media. Furthermore, novel capping and tagging methods[15] developed for automated synthesis help to greatly simplify the post-synthetic workup and purification process of synthetic oligosaccharides. Following each coupling step, unreacted hydroxyl groups that may give rise to shorter carbohydrate sequences are treated with a capping reagent that renders them silent in subsequent couplings.

Traditionally, branched carbohydrates such as the Lewis antigens have been synthesized in solution by highly convergent routes.[16,17] The Lewis X pentasaccharide, the Lewis Y hexasaccharide, and dimeric combinations of Lewis antigens, including the Le^y–Le^x nonasaccharide, are blood group determinant oligosaccharides. The last two also act as tumor markers and are currently being explored in cancer therapy.[18] With our sequential strategy using a small number of glycosyl donors **9**–**13** as building blocks an automated solid-phase synthesis of fully protected Lewis blood group oligosaccharides **6**–**8** was possible (Scheme 2).[12]

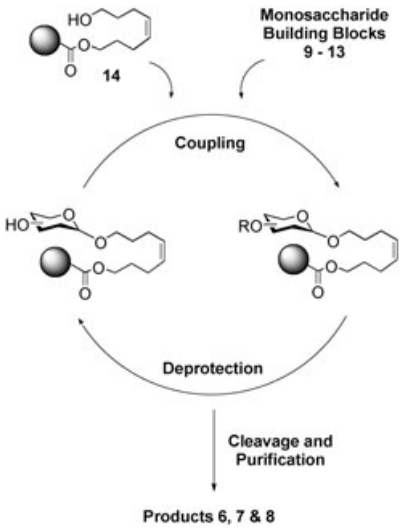


Scheme 2. Retrosynthesis of the protected Lewis X pentasaccharide **6**, Lewis Y hexasaccharide **7**, and Le^x-Le^y nonasaccharide **8** indicates monosaccharide building blocks **9**–**13**. Bn = benzyl, Bu = butyl, Fmoc = 9-fluorenylmethoxycarbonyl, Lev = levulinoyl, Piv = pivaloyl, TCA = trichloroacetyl.

Activation of the glycosyl phosphate monomers **9**–**13** was carried out at −15 °C in dichloromethane under acidic conditions with the Lewis acid TMSOTf (TMSOTf = trimethylsilyl trifluoromethanesulfonate). Removal of Fmoc was accomplished by treatment with excess piperidine, whereas the levulinoyl group was removed reductively by treatment with a solution of hydrazine. The use of UV-active protective groups such as Fmoc is a great advantage and allows for the real-time monitoring of the success of the automated synthesis as is common for the synthesis of peptides and oligonucleotides. All the coupling and deprotection steps were repeated at least twice. A general cycle for the incorporation of one building block is shown in Table 1. Repetition of these cycles (Scheme 3) with the corresponding building

Table 1. General cycle used with glycosyl phosphates for the construction of oligosaccharides **6**–**8**.

Step	Function	Reagent	*t* [min]
1	couple	5 equiv donor and 5 equiv TMSOTf	21
2	wash	dichloromethane	9
3	couple	5 equiv donor and 5 equiv TMSOTf	21
4	wash	*N*,*N*-dimethylformamide (DMF)	9
5	deprotection	3 × 175 equiv piperidine in DMF	34
		or 5 × 10 equiv hydrazine in DMF	80
6	wash	*N*,*N*-dimethylformamide (DMF)	9
7	wash	0.2 M acetic acid in tetrahydrofuran	9
8	wash	tetrahydrofuran	9
9	wash	dichloromethane	9



Scheme 3. Automated oligosaccharide synthesis with glycosyl phosphates. Initial glycosylation of resin-bound acceptor **14** produces a coupling product that is subsequently deprotected. Iteration of coupling and deprotection cycles with phosphate donors **9**–**13** followed by cleavage of the resin-bound oligosaccharides and purification furnishes **6**–**8**.

blocks completed the assembly of the penta-, hexa-, and nonasaccharide in less than one day.

Cleavage of the ester linker from the resin using a solution of sodium methoxide over a period of six hours provided the crude oligosaccharides. HPLC purification produced the fully protected Lewis X pentasaccharide (**7**) in 12.6% yield, Lewis Y hexasaccharide (**8**) in 9.9% yield, and Le^y–Le^x nonasaccharide (**9**) in 6.5% yield.[12]

Mapping Interactions of Carbohydrates in Biological Systems

Once a carbohydrate structure of biological interest has been synthesized, several tools[19] to map interactions of the carbohydrates in biological systems are at the disposal of today's glycobiologist. These tools include modified surfaces for microarrays and surface plasmon resonance (SPR) spectroscopy, monovalent fluorescent conjugates, neoglycoproteins and carbohydrate vaccines, multivalent quantum dot conjugates, affinity-tagged saccharides, derivatized magnetic particles, and latex microspheres (Figure 1). All these methods rely on clever linking chemistries. Amine-containing linkers are able to react with amine-reactive substrates such as activated esters. In analogy, carboxy-group-containing linkers react with amine-containing molecules. Furthermore, thiol-containing linkers react readily with maleimide and iodoacetyl moieties and vice versa. In addition, thiol-containing moieties show a high affinity to gold surfaces.

One special linker has been devised for the majority of the tools. 2-[2-(2-Mercaptoethoxy)ethoxy]ethanol was selected due its compatibility with existing synthetic methods, the ease of temporarily masking the thiol functionality with a protecting group, and the readily applicable thiol-based conjugation chemistry (Scheme 4).

Carbohydrate microarrays: Microarrays[20] in the "chip" format, prepared by attachment of biopolymers to a surface in a spatially discrete pattern, have enabled a low-cost and high-throughput methodology for screening interactions involving these molecules. Compared to classical methods, microarrays allow for several thousand binding events to be screened in parallel, whereby one experiment requires only miniscule amounts of both analyte and ligand. Thus, binding profiles and lead structures are readily examined. Miniaturization through the construction of microarrays is particularly well suited for investigations in the field of glycomics.[21] In contrast to the other two classes of biopolymers, no biological amplification strategy, such as the polymerase chain reaction (PCR) or cloning, exists to produce usable quantities of complex oligosaccharides. Therefore, the miniaturized

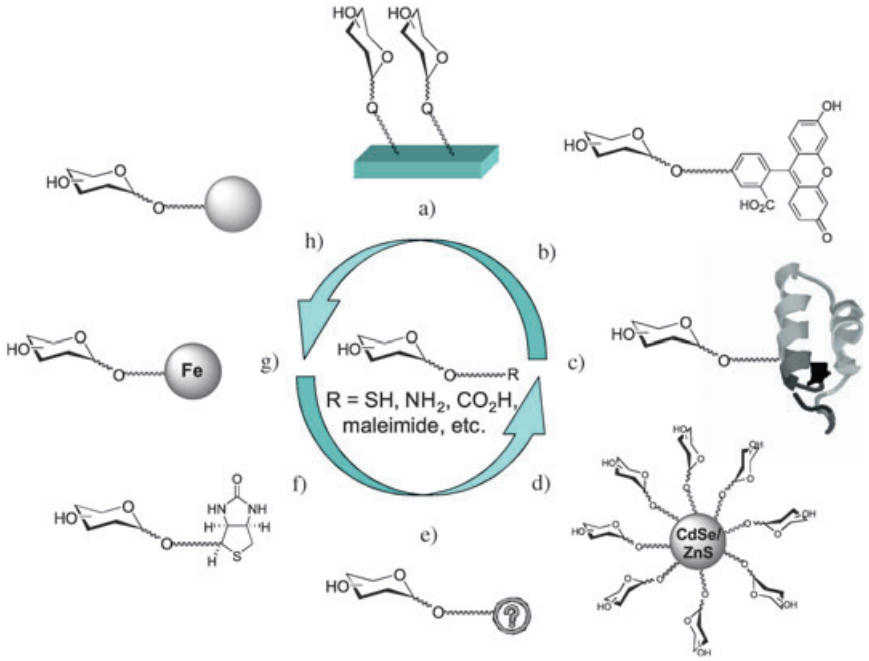
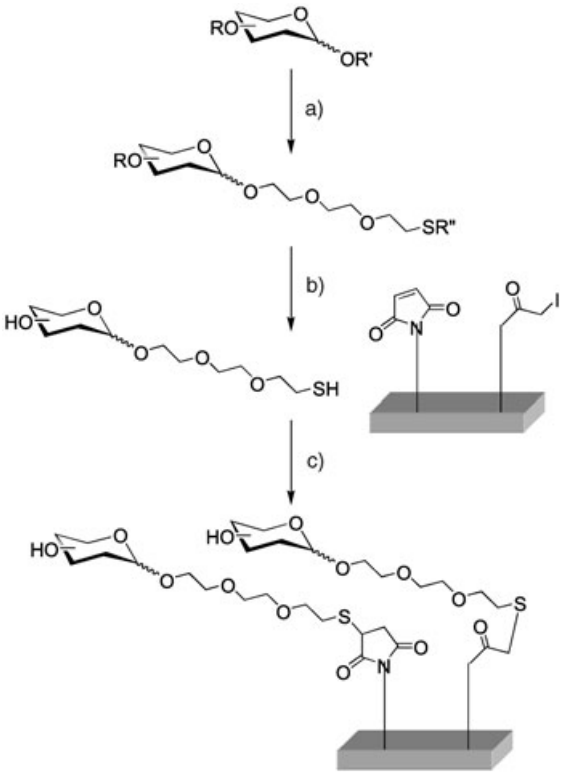


Figure 1. Tools for glycobiology: a) modified surfaces for microarrays and surface plasmon resonance (SPR); b) monovalent fluorescent conjugates; c) neoglycoproteins and carbohydrate vaccines; d) multivalent quantum dot conjugates; e) future neoglycoconjugates; f) affinity tag conjugates; g) magnetic particle conjugates; h) latex microsphere and sepharose affinity resin conjugates.



Scheme 4. 2-[2-(2-Mercaptoethoxy)ethoxy]ethanol as a linker for preparing neoglycoconjugates: a) Linker synthetically incorporated into reducing end of mono- or oligosaccharide. b) All protecting groups are removed from carbohydrate and thiol. c) Reduced thiol coupled to maleimide or iodoacetyl functionalized structure (chip, bead, resin, fluorescent dye, quantum dot, etc.).

assay format is the method of choice to perform several experiments with only 10^{-12} mol of compound.

Hitherto, many methods for the preparation of carbohydrate microarrays have been described, such as nitrocellulose coated slides for noncovalent immobilization of microbial polysaccharides[22] and self-assembled monolayers modified by Diels–Alder mediated coupling of cyclopentadiene-derivatized oligosaccharides,[23] just to name two. Unfortunately, the first method requires large polysaccharides or lipid-modified sugars for the noncovalent interaction. The latter method requires the preparation of oligosaccharides bearing the sensitive cyclopentadiene moiety.

In our laboratory, the best results were obtained by utilizing maleimide functionalization of glass slides and the immobilization of the oligosaccharides with thiol-containing linkers. However, with this linker system two methods of surface functionalization should be distinguished: One presents a relatively low density of immobilized oligosaccharides and excellent resistance to nonspecific binding of proteins to the chip surface. The other permits a high-density immobilization of carbohydrates, and thus, allows for the examination of oligosaccharide clusters at the surface. In this case, the carbohydrate is presented in a peptide-free context.

Hybrid carbohydrate/glycoprotein microarrays: A chip containing both carbohydrates and glycoproteins permits the rapid determination of the context of binding to the glycoprotein. Incubation of proteins with this hybrid array establishes whether the peptide context is essential for binding or whether the carbohydrate structure alone is sufficient. To prepare these slides, the glass surface is usually modified with two different chemistries, for example, on one side a maleimide chemistry, and on the other an *N*-hydroxysuccinimide (NHS) activated ester.

Microsphere arrays: In contrast to common microarrays, the microsphere system uses optical methods to define the position and structure of a carbohydrate series.[24] Incubation of the immobilized microsphere with a fluorophore-labeled carbohydrate binding protein and the subsequent measuring of the fluorescence signals permits a determination of the binding profile. Binding events take place when one bead emits at both the wavelength of an internal code, which is used as marker for the oligosaccharide, attached to the microsphere, and the fluorophore-labeled protein.

Surface plasmon resonance (SPR) spectroscopy: A method to get quantitative insights into binding of analytes to ligands in real time is surface plasmon resonance (SPR).[25] For SPR experiments, one of the interacting species is immobilized on the surface of a chip. The prospective binding partner is allowed to flow over the chip. During this process the refractive index of the chip changes due to the interaction as well as the accumulation of analyte. The kinetic data obtained in this fashion allows one to calculate association and dissociation constants from submicrogram quantities of material. There is no need to label the ligand or analyte, and any influence of a label on the binding affinities can be excluded. A further advantage is that these measurements permit evaluations of low- and high-affinity interactions. SPR is on the way to becoming an extremely powerful tool in glycomics, since structure–activity relationships are quickly assessed.

Fluorescent carbohydrate conjugates: Microarrays do not represent ideal formats for the examination of monovalent protein–carbohydrate interactions. Commonly, the densities of the immobilized oligosaccharides are too high to ensure that monovalent interactions are observed. Another limitation of the array technique is the requirement of purified receptor. Therefore, another, more appropriate approach is needed to study interactions with cells.

Monovalent and multivalent fluorescent probes can be utilized to evaluate the influence of oligosaccharide clustering on recognition by cell-surface lectins. Fluorescence microscopy and flow cytometry are appropriate methods to visualize the corresponding receptor–carbohydrate interactions.

Carbohydrate affinity screening: In contrast to the array technique that usually utilizes purified receptors, this syn-

thetic tool facilitates the isolation and purification process of carbohydrate-binding proteins.[19] Crude mixtures or biological extracts are separated by carbohydrate-containing affinity columns. Thus, this purification method provides also information about the interaction of carbohydrates with other biopolymers.

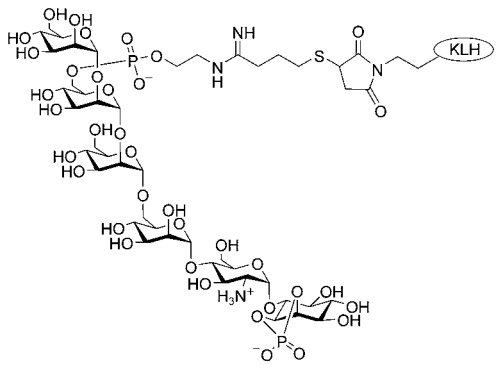
Oligosaccharide Conjugate Vaccines: Malaria and HIV

In addition to serving as tools, carbohydrates also hold great potential as vaccines, as small amounts of antigen can be used to protect a large number of people. Immunological investigations that utilize fully synthetic carbohydrate vaccines have shown very promising results in the treatment of various diseases. These affiliations include cancer, bacterial infections, such as tuberculosis, and tropical diseases, such as leishmaniasis and malaria.

The malaria parasite *Plasmodium falciparum*, which infects 5–10% of humanity, accounts for about 100 million clinical cases and the deaths of more than two million people annually due to inflammation caused by the malaria toxin.[26] Current malaria treatments are often impractical, and drug resistance is a growing problem. At the same time, there is still no effective malaria vaccine.

The malaria parasite expresses a large amount of glycosylphosphatidylinositol (GPI) on the cell surface. There is evidence that the inflammatory cascade triggerd by this GPI is responsible for much of malaria's morbidity and mortality.

To prepare this antigen, the synthetic hexasaccharide malaria toxin **15**[27] was treated with a linker and conjugated to a maleimide-activated carrier protein. Mice that were treated with chemically synthesized GPI attached to the protein were substantially protected from death by malaria. Between 60 and 75% of the vaccinated mice survived, whereas the survival rate for unvaccinated mice was only 0–9%. It should be noted that only miniscule amounts (10^{-9}–10^{-7} g per person) of the hexasaccharide **15** that was partly assembled by automated synthesis will be necessary to perform the vaccination. The preclinical model revealed that a nontoxic GPI oligosaccharide coupled to a carrier protein is immunogenic and provides significant protection against malarial pathogenesis. An antitoxic oligosaccharide vaccine against malaria might be within reach.

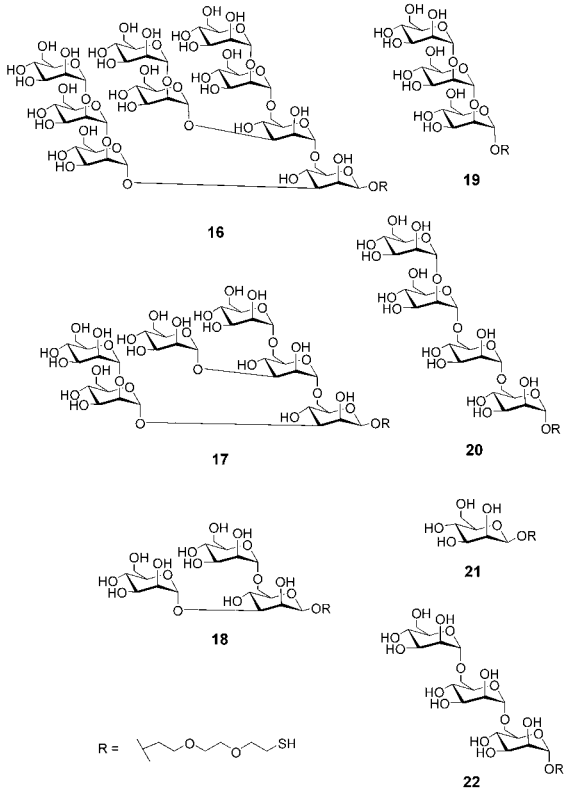


15

The elucidation of HIV envelope glycoprotein interactions with prospective binding partners advances our understanding of the viral entry and provides a basis for the design of new vaccines interfering with HIV entry. Using the chip format, interactions of carbohydrates decorating the viral surface envelope proteins with receptors are readily discovered. Relevant substructures that are important for binding can be identified simultaneously when the arrays are composed of a series of closely related analogues.[28]

One important carbohydrate structure found in the HIV envelope glycoprotein gp120 is the triantennary *N*-linked mannoside $(Man)_9(GlcNAc)_2$. Utilizing a variety of synthetic mannose-containing substructures **16**–**22**, a chip with a wide range of concentrations was printed in order to establish a saturation point for observed binding to a fluorescently labeled protein.[28] Thus, a carbohydrate-binding profile can be established for a given protein by comparing the integrated fluorescence between different spots.

Incubation of these arrays with a series of different gp120 binding proteins (ConA, 2G12, Cyanovirin-N, DC-SIGN, and Scytovirin-N) revealed a precise evaluation of their binding profiles.[28] Figure 2 shows the corresponding chips. The experiments with 2G12 showed no binding to **18**, **21**, and **22** suggesting that a Manα1–2Man linkage, the only structural motif in common, is necessary for recognition by 2G12. In contrast, Scytovirin-N, a protein that was isolated from the cyanobacterium *Scytonema varium*, binds only to



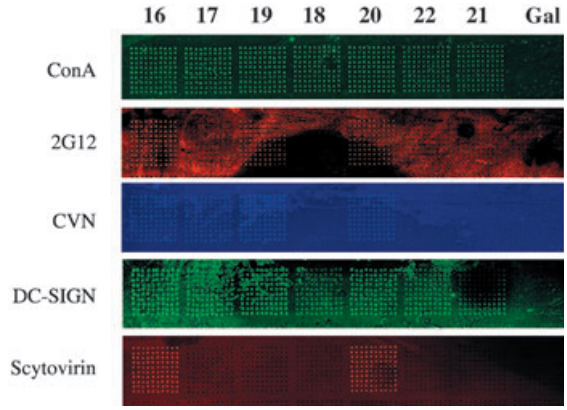


Figure 2. Carbohydrate microarrys containing synthetic mannans **16–22** and galactose, printed at 2 mM. Each carbohydrate is spotted with a diameter of approximately 100–200 μm. False color image of incubations with fluorescently labeled ConA, 2G12, CVN, DC-SIGN, and Scytovirin.

the structures **16** and **20**. This result clearly illustrates that a different structural motif within the oligosaccharide is recognized by Scytovirin-N. The terminal Manα1–2Man linkage, together with the underlying α1–6 trimannoside moiety is necessary for Scytovirin-N binding. These studies also corroborate that these proteins can bind high-density arrays of Manα1–2Man-containing oligosaccharides in the absence of the polypeptide backbone.

Carbohydrate–Nucleic-Acid Interactions: Aminoglycosides

Aminoglycosides present a family of natural-occuring pseudooligosaccharides that consist of two to five monomers and contain almost a one-to-one ratio between amino and hydroxy groups. Clinically, these compounds have been used to treat infectious diseases induced by a variety of Gram-negative bacteria. Aminoglycosides exhibit their antibiotic activity by inhibiting protein synthesis by binding to bacterial ribosomes. Most commonly, aminoglycosides bind to the A site in the small ribosomal subunit (30S) of the bacterial ribosome resulting in misreading during the translational process. Not surprisingly, charge interactions between amino groups and the phosphate backbone dominate as binding forces in these aminoglycoside–RNA complexes. As with many other antibiotics, the efficiency of aminoglycosides has been compromised by the emergence of resistant bacterial strains.[29,30] The most prominent mechanisms that cause resistance are enzymatic modifications of the aminoglycoside including N-acetylation and O-phosphorylation. These modifications result in a large decrease in binding affinity to the therapeutic target.[31]

To facilitate the discovery of safer and more active aminoglycosides, high-throughput methods are needed. Microarray techniques enable medicinal chemists to identify weak binders to resistance-causing enzymes and tight binders to ribosomal RNA. Recently, we reported the construction of aminoglycoside microarrays to study antibiotic resistance.[32,33] The antibiotics were immobilized on amine reactive glass slides by using a DNA-arraying robot. Two aminoglycoside acetyltransferase resistance enzymes, 2′-acetyltransferase (AAC(2′)) from *Mycobacterium tuberculosis*[34] and 6′-acetyltransferase (AAC(6′)) from *Salmonella enterica*[35] were used as examples. Hybridization to the aminoglycoside arrays revealed that each aminoglycoside interacts with both enzymes. Comparison with calorimetric studies of aminoglycoside binding affinities to AAC(6′)[36] indicated a strong correlation with the array results. Arrays were also incubated with two different RNA sequences in order to determine binding specifity for bacterial and human A-site RNA.

To facilitate the discovery of inhibitors of resistance-causing enzymes, a library of aminoglycoside mimetics was synthesized and immobilized on arrays. Guanidinoglycosides[37] (Figure 3) were chosen as aminoglycoside analogues for several reasons:

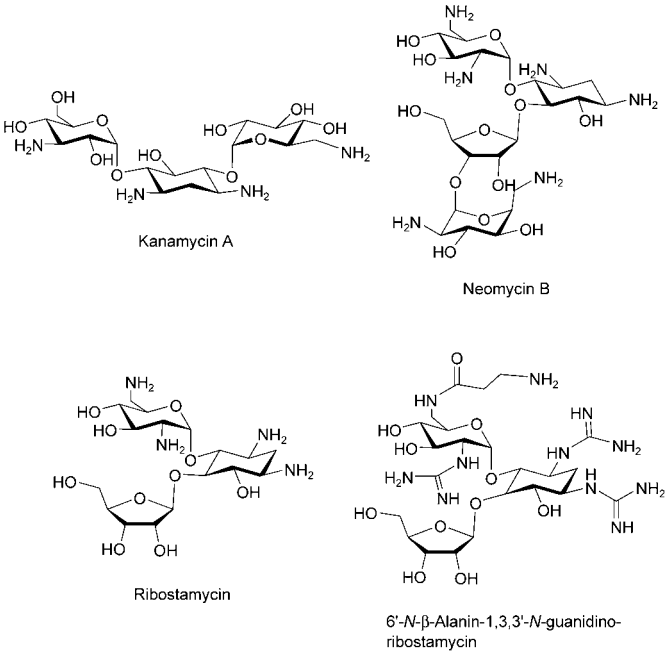


Figure 3. Representative examples of aminoglycosides (Kanamycin A, Neomycin B, Ribostamycin). Furthermore, a guanidinoglycoside (6′-*N*-β-alanin-1,3,3′-*N*-guanidinoribostamycin) with a corresponding linker for immobilization chemistry is shown.

1) Guanidinoglycosides can readily prepared from aminoglycosides.
2) The increased positive charge due to the larger number of nitrogen-containing guanidino groups may allow guanidinoglycosides to bind more tightly to the negatively charged aminoglycoside binding pocket.[38]
3) The large difference in the pK_a values of guanidino and amino groups (12.5 versus 8.8) suggests that guanidinoglycosides are likely not substrates for acetyltransferases such as AAC(2′) and AAC(6′).

As anticipated, guanidinoglycosides revealed higher affinity to resistance-causing enzymes than the corresponding aminoglycosides. Guanidinoglycosides do not serve as substrates and inhibit acylation of several clinically important antibiotics. This promising approach proves valuable for screening a plethora of compounds in a short time to discover improved drugs that evade current modes of bacterial resistance.

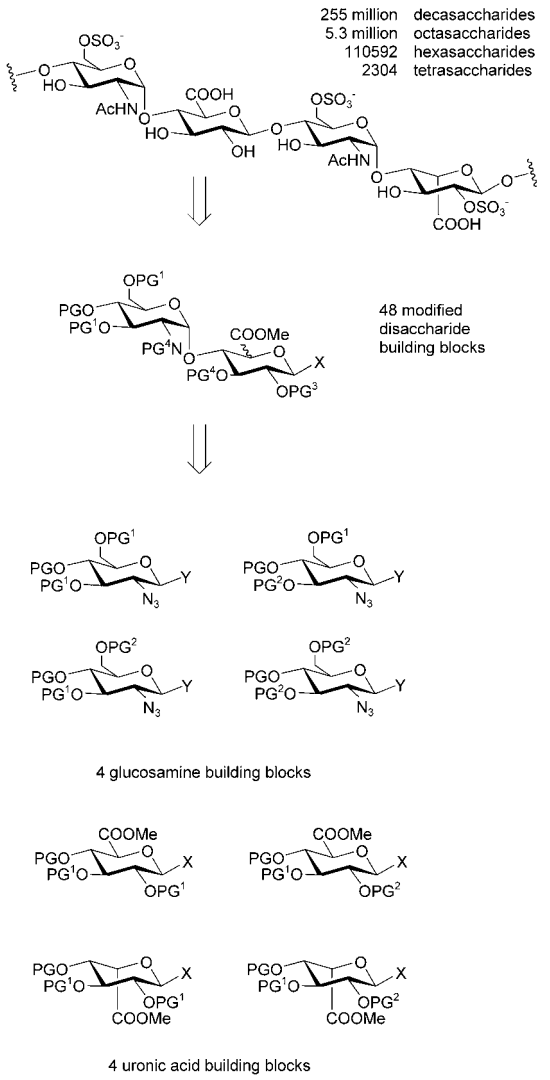
Carbohydrate–Protein Interactions: Heparin

Heparin is a biologically important and chemically unique polysaccharide and a key player in a plethora of different physiological processes.[39] The interaction with the protein antithrombin III (AT III) is best understood. Thus, since the late 1930s, heparin has served as a clinical anticoagulant in the treatment of heart disease. Interactions with growth factors, chemokines, lipid-binding proteins, and viral-envelope proteins are of interest.[39]

Heparin is a linear, unbranched, highly sulfonated polymer that consists of 1→4-linked pyranosyluronic acid and glucosamine units (Scheme 5).[40] The type of uronic acid building block varies; usually 90% of L-iduronic acid and 10% of D-glucuronic acid are found. Commonly, 20 to 200 disaccharide repeat units are found, giving rise to tremendous structural complexity. Over the last two decades, a variety of different synthetic methods to assemble heparin oligosaccharides have been disclosed. Instead of developing a new total synthesis of each heparin-like glycosaminoglycan sequence, modular, highly convergent synthetic approaches for the rapid assembly of defined heparin oligosaccharide sequences were developed.[41,42] Such a synthetic strategy employs closely related building blocks (modules), whose similarity should allow for standardized condensation reactions and provide flexibility and variety in use. Careful consideration of many synthetic challenges is required due to the great diversity of the native structures. The sulfation patterns mandate the placement of specific protecting groups in all positions to carry sulfates and different protection on hydroxyl groups that remain unaltered. Furthermore, a protecting-group scheme to differentiate the glucosamine amine group is required, because the amine is found to be unaltered, acetylated, or sulfonated.[43] In total, 48 modified disaccharide building blocks are needed (Scheme 5).

Several key issues had to be addressed for us to be able to apply our modular approach based on the assembly of disaccharide modules: 1) A pentenol moiety that allows further functionalization and mimics the situation on solid support was used. 2) Uronic acid glycosyl donors were employed to avoid difficult late stage oxidations. 3) The protecting-group pattern that was chosen is similar to that of previous total syntheses. Thus, well-established deprotection and sulfation protocols could be used to generate the final oligosaccharide structure.

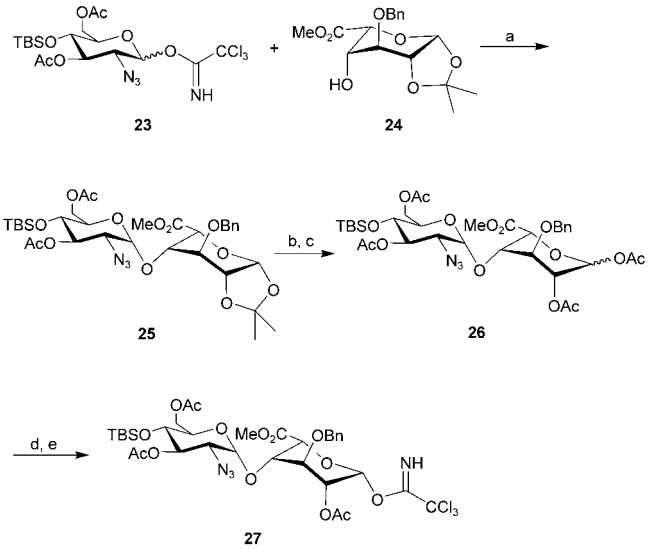
Below, a representative example for the modular assembly of a heparin tetrasaccharide is described.[42] A novel



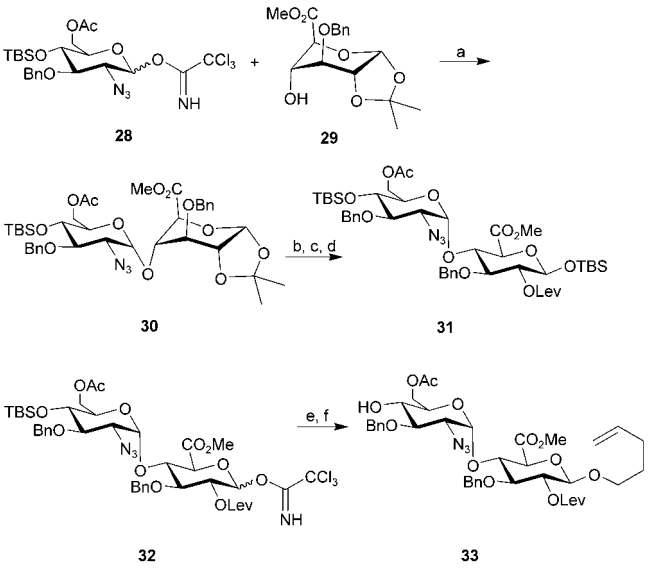
Scheme 5. Retrosynthetic analysis of a general, modular approach to the preparation of heparin-like glycosaminoglycans (PG = protecting group).

method was developed for the construction of the disaccharide building blocks with the α-D-glucosazide-D-uronic acid linkage. The uronic acid acceptors **24** and **29** were locked in a $^1C_4$-conformation (Schemes 6 and 7). Having constructed a collection of both glucosazido–iduronic acid and glucosazido–glucuronic acid dimer building blocks (in Schemes 6–8 only one example is depicted), the modular synthesis can be carried out.

The elongation of the heparin oligosaccharide sequence was performed by combining disaccharide trichloroacetimidate **27** with the reducing-end module **33**. After having assembled the core structure of the heparin tetrasaccharide **34**, deprotection and sulfation had to be performed. Acetate groups mark positions to be sulfated and benzyl ether groups designate free hydroxyl moieties. After the removal of acetate groups under basic conditions, the free hydroxyl groups can be sulfated by Et_3NSO_3. Cleavage of all benzyl ether moieties and the reduction of the azide group were



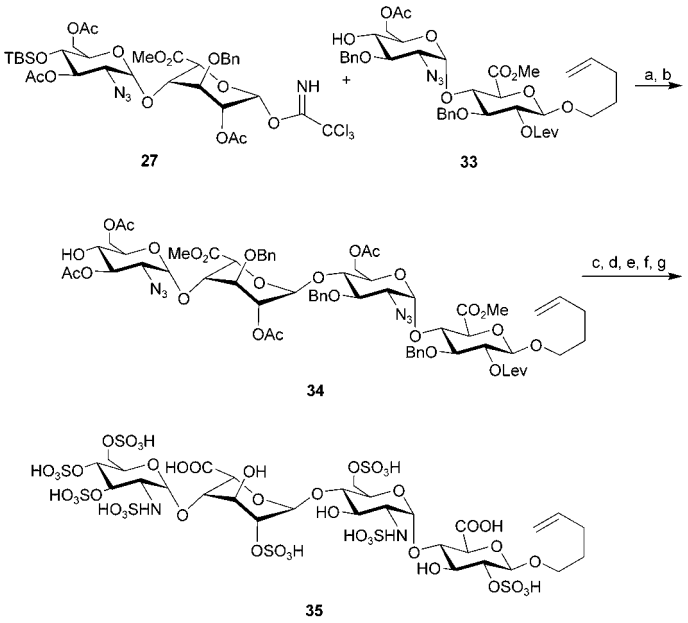
Scheme 6. Synthesis of iduronic acid disaccharide building block **27**. a) TBSOTf, 4 Å molecular sieves, CH_2Cl_2, −30 °C→room temperature (90 %); b) 60 % aq dichloroacetic acid (89 %); c) Ac_2O, CH_2Cl_2, DMAP, pyridine (95 %); d) $BnNH_2$; Et_2O, 0 °C; e) CCl_3CN, DBU, CH_2Cl_2, 0 °C (67 %, 2 steps). DBU = 1,8-diazabicyclo[5.4.0]undec-7-ene, DMAP = 4-dimethylamino pyridine, TBSOTf = *tert*-butyldimethylsilyl triflate.



Scheme 7. Synthesis of reducing end disaccharide building block **33**. a) TBSOTf, 4 Å molecular sieves, CH_2Cl_2, −78 °C→room temperature (86 %); b) 50 % aq dichloroacetic acid (81 %); c) TBSCl, imidazole, CH_2Cl_2, 0 °C; d) $(Lev)_2O$, DMAP, CH_2Cl_2 (81 %, 2 steps); e) TBAF, HOAc, THF; f) CCl_3CN, DBU, CH_2Cl_2, 0 °C (92 %, 2 steps). TBSCl = *tert*-butyldimethylsilyl chloride, TBAF = tetrabutylammonium fluoride, $(Lev)_2O$ = levulinic anhydride.

performed by Pd-catalyzed hydrogenation. Selective N-sulfation furnished fully functionalized heparin tetrasaccharide **35**.[42]

Due to the high content of negatively charged sulfate and carboxyl groups, the most prominent type of interaction between heparin and basic amino acids of proteins is of ionic



Scheme 8. Coupling to the tetrasaccharide **34**, protecting group modifications, sulfation, and final deprotection to afford **35**. a) TMSOTf, CH_2Cl_2, −25 °C (88 %); b) HF/pyridine, AcOH, THF (86 %); c) $NH_2NH_2{\cdot}H_2O$, pyridine, AcOH (90 %); d) 1. LiOH (0.7 M aq), H_2O_2 (50 % aq), THF, overnight; 2. 4 M NaOH, room temperature overnight (82 %, 2 steps); e) Et_3NSO_3, DMF, 50 °C, overnight (50 %); f) H_2, Pd/C, EtOH, water (quant); g) SO_3·pyridine, H_2O (60 %).

nature. In some cases, also hydrogen bonding and even hydrophobic interactions are not negligible. With the exception of the AT-III–heparin interaction, for which the exact sequence of heparin associating with the protein has been identified, the structure–function relationship of heparin is still quite poorly understood. A better understanding is necessary to apply defined heparin sequences to the treatment of other diseases. A variety of techniques including surface plasmon resonance (SPR) have been applied to study heparin–protein interactions.[39]

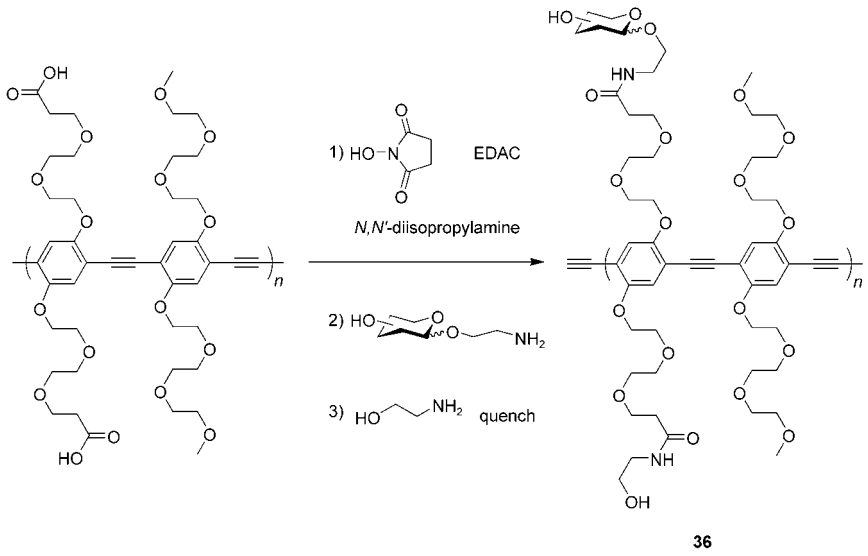
Detection of Pathogenic Bacteria

Usually, the detection of pathogenic bacteria, such as *Escherichia coli*, is based on the selective growth of these bacteria in liquid media or on plates. This procedure may require several days.[44] More recently, methods such as pathogen recognition by fluorescently labeled antibodies, DNA probes, or bacteriophages have been developed and proved to be much faster.[44]

In many cases, bacteria as well as viruses bind to carbohydrates displayed on the host cells they infect. *Escherichia coli* binds mannose influenza virus binds to sialic acid, to name two examples.[45] To ensure the high binding affinity necessary for strong adhesion and successful infection of the cell, the pathogen often uses multivalent interactions.[46] Conducting polymers displaying carbohydrates can simulate

these binding events and serve as an ideal material to detect even small amounts of pathogens.

Recently, we reported a carbohydrate-functionalized poly(*p*-phenylene ethynylene) (PPE) **36** that can be used for the detection of *Escherichia coli* by multivalent interactions (Scheme 9).[47] Therefore, 2′-aminoethyl mannoside and galactoside were coupled to PPE. Unreacted succinimide esters were quenched by addition of excess ethanolamine, before washing with water to remove uncoupled reagents. The loading of the polymer was determined by a phenol sulphuric acid test and revealed that about 25% of the reactive sites were functionalized with glycosides. A fluorescence resonance energy transfer (FRET) experiment insured that mannose-binding lectin moieties interact with mannose displayed on the polymer without affecting binding selectivity and do not exhibit any nonspecific binding. Experiments with two bacterial strains differing in their mannose-binding properties revealed that the mannose-functionalized polymer imparted strong fluorescence to mannose-binding *Escherichia coli*. Even separation and rinsing procedures are not able to remove the bacteria from the polymer. In contrast, the mutated strain unable to bind mannose showed no signal and no aggregation of bacteria.



36

Scheme 9. Synthesis of the carbohydrate-functionalized fluorescent polymer **36** for the detection of pathogenic bacteria.

The binding events involving the functionalized polymers and the bacteria were followed with the microscope. Mutant bacteria that lost the ability to bind to mannose did not bind to the polymer, whereas the mannose-binding bacteria aggregated in clusters with fluorescent centers (Figure 4). The number of cells in these clusters varies between 30 and several thousand. As anticipated, the larger the aggregates, the stronger the fluorescence signal. Competitive binding experiments with other carbohydrates displayed on the polymer do not reveal any fluorescent clusters. To determine the detection limit of this new method, serially diluted solutions of mannose-binding *E. coli* were incubated with the mannose-containing polymer. Fluorescence microscopy experiments revealed a limit in the range of 10^3–10^4 bacteria. Similar values were obtained before by using fluorescently labeled antibodies. Further competition experiments have shown that only relatively high concentrations of free mannose (10 mM) inhibit binding to the polymer significantly. At concentrations of less than 10 μM the clustering is not affected at all.

However, many pathogens bind the same carbohydrates, for example, *E. coli* as well as *Salmonella enterica* bind to mannose. This limitation may be overcome using multiple carbohydrates on arrays.[48]

Conclusion and Outlook

The synthesis, isolation, purification, and structure elucidation of carbohydrates has been a challenging goal for decades. Recently, new methods to gain

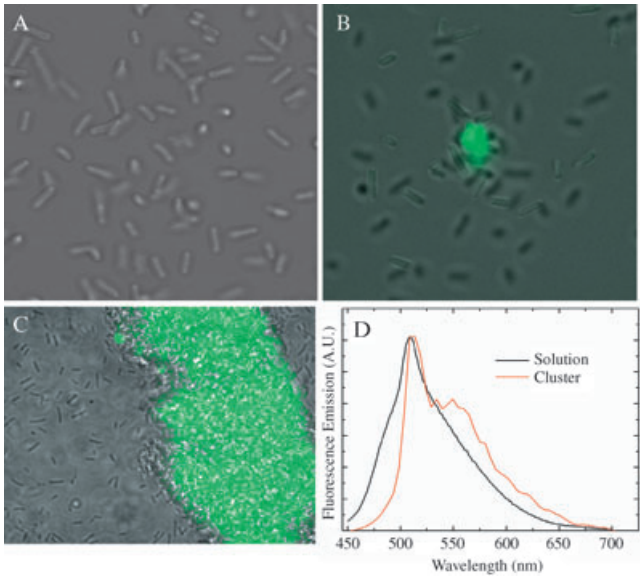


Figure 4. A) Laser-scanning confocal microscopy image of mutant *Escherichia coli* that does not bind to polymer **36**. B) A fluorescent bacterial aggregate forms due to multivalent interactions between the mannose-binding bacterial pili and the polymer **36** (superimposed fluorescence and transmitted light images). C) Fluorescence microscopy image of a large fluorescent bacterial cluster. D) Conventional fluorescence spectra of polymer **36** (black) and normalized fluorescence spectra of a bacterial cluster obtained using confocal microscopy (red).

access to these complex molecules have been developed, including a fully automated oligosaccharide synthesizer. Easily accessible glycosyl phosphates and glycosyl trichloroacetimidates proved to be a powerful class of glycosylating agents for this purpose. High-yielding coupling steps on the solid support rely on the use of an excess amount of building blocks in the presence of a stoichiometric amount of TMSOTf. Suitable protection and deprotection strategies lead to the assembly of linear and even branched oligosaccharides that can now be performed in a fully automated manner.

Several tools to understand the intricate role of oligosaccharides in various cell-signaling processes have been developed. The "chip" format enables glycoscientists to elucidate interactions of carbohydrates with fluorescently labeled proteins, including bacterial and viral toxins. Clever linking chemistries provide a wider range of glycans available for screening in the microarray format. The chips are constructed by using standard DNA gene chip instrumentation. To detect interactions, only miniscule amounts of both ligand and analyte are necessary.

The tool kit consisting of carbohydrate synthesizer and carbohydrate microarrays lays the foundation for the discovery and elucidation of new drugs, as studies with the fully synthetic antitoxin malaria vaccine have shown. HIV-neutralizing proteins have been identified by studies with carbohydrate microarrays; aminoglycoside microarrays were used to test antibacterial resistance. Fluorescent polymers can be utilized to detect small amounts of pathogenic bacteria in a short time.

Whereas many complex carbohydrate structures of pyranosides are now accessible by automated synthesis, the automated assembly of bacterial sugars is still a difficult goal to achieve. A further bottleneck is a rapid and highly efficient synthesis of the monosaccharide building blocks. More efficient syntheses for most of the approximately 50 carbohydrate building blocks are required.

Future glycobiologists will be able to screen a plethora of complex carbohydrates that are thought to play previously unimaginable roles in biological systems. The knowledge gained from glycomics will be as important as a basis for the pharmaceutical industry as that discovered in the field of genomics and proteomics during the last 30 years. We are still just beginning to understand the importance of carbohydrates in biological information transfer and storage. Automation in the construction of glycosidic linkages will allow chemists and biologists to spend less time asking the question "How do I make this structure?" and more time answering the question: "What does this structure do?".

Acknowledgements

We thank all present and past members of the Seeberger group and our collaborators who contributed to the results reported in this chapter. D.B.W. is grateful to the Alexander von Humboldt Foundation for a Feodor Lynen Research Fellowship and to the Deutsche Forschungsgemeinschaft (DFG) for a Emmy Noether Fellowship. P.H.S. thanks the ETH for financial support, and for grant SNF 200021-101593/1.

[1] a) P. Nissen, J. Hansen, N. Ban, T. A. Steitz, *Science* **2000**, *289*, 920–930; b) N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 905–920.
[2] a) A. Varki, *Glycobiology* **1993**, *3*, 97–130; b) H. Lis, N. Sharon, *Eur. J. Biochem.* **993**, *218*, 1–27; c) R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683–720; d) Y. C. Lee, R. T. Lee, *Acc. Chem. Res.* **1995**, *28*, 322–327; e) W. H. Chambers, C. S. Brisette-Storkus, *Chem. Biol.* **1995**, *2*, 429–435.
[3] a) M. H. Caruthers, *Science* **1985**, *230*, 281–285; b) M. H. Caruthers, *Acc. Chem. Res.* **1991**, *24*, 278–284.
[4] E. Atherton, R. C. Sheppard, *Solid-Phase Peptide Synthesis: A Practical Approach*, Oxford University Press, Oxford, **1989**.
[5] a) R. Rodebaugh, S. Joshi, B. Fraser-Reid, H. M. Geysen, *J. Org. Chem.* **1997**, *62*, 5660–5661; b) J. Rademann, A. Geyer, R. R. Schmidt, *Angew. Chem.* **1998**, *110*, 1309–1313; *Angew. Chem. Int. Ed.* **1998**, *37*, 1241–1245.
[6] a) S. J. Danishefsky, M. T. Bilodeau, *Angew. Chem.* **1996**, *108*, 1482–1522; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1380–1419; b) P. H. Seeberger, S. J. Danishefsky, *Acc. Chem. Res.* **1998**, *31*, 685–695; c) R. R. Schmidt, J. C. Castro-Palomino, O. Retz, *Pure Appl. Chem.* **1999**, *71*, 729–744.
[7] C.-H. Wong, *Pure Appl. Chem.* **1995**, *67*, 1609–1616.
[8] O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science* **2001**, *291*, 1523–1527.
[9] P. H. Seeberger, *Chem. Commun.* **2003**, 1115–1121.
[10] O. J. Plante, R. B. Andrade, P. H. Seeberger, *Org. Lett.* **1999**, *1*, 211–214.
[11] R. R. Schmidt, W. Kinzy, *Adv. Carbohydr. Chem. Biochem.* **1994**, *50*, 21–123.
[12] K. R. Love, P. H. Seeberger, *Angew. Chem.* **2004**, *116*, 612–615; *Angew. Chem. Int. Ed.* **2004**, *43*, 602–605.
[13] M. C. Hewitt, P. H. Seeberger, *Org. Lett.* **2001**, *3*, 3699–3702.
[14] R. B. Andrade, O. J. Plante, L. G. Melean, P. H. Seeberger, *Org. Lett.* **1999**, *1*, 1811–1814.
[15] E. R. Palmacci, M. C. Hewitt, P. H. Seeberger, *Angew. Chem.* **2001**, *113*, 4565–4569; *Angew. Chem. Int. Ed.* **2001**, *40*, 4433–4437.
[16] G. Hummel, R. R. Schmidt, *Tetrahedron Lett.* **1997**, *38*, 1173–1176.
[17] P. P. Deshpande, S. J. Danishefsky, *Nature* **1997**, *387*, 164–166.
[18] G. Ragupathi, P. P. Deshpande, D. M. Coltart, H. M. Kim, L. J. Williams, S. J. Danishefsky, P. O. Livingston, *Int. J. Cancer* **2002**, *99*, 207–212.
[19] D. M. Ratner, E. W. Adams, M. D. Disney, P. H. Seeberger, *ChemBioChem* **2004**, *5*, 1–10.
[20] a) D. Barnes-Seemann, S. B. Park, A. N. Koehler, S. L. Schreiber, *Angew. Chem.* **2003**, *115*, 2478–2481; *Angew. Chem. Int. Ed.* **2003**, *42*, 2376–2379; b) S. Fukui, T. Feizi, C. Galustian, A. M. Lawson, W. Chai, *Nat. Biotechnol.* **2002**, *20*, 1011–1017; c) A. N. Koehler, A. F. Shamji, S. L. Schreiber, *J. Am. Chem. Soc.* **2003**, *125*, 8420–8421; d) P. J. Hergenrother, K. M. Depew, S. L. Schreiber, *J. Am. Chem. Soc.* **2000**, *122*, 7849–7850.
[21] a) S. Bidlingmaier, M. Snyder, *Chem. Biol.* **2002**, *9*, 400–401; b) K. R. Love, P. H. Seeberger, *Angew. Chem.* **2002**, *114*, 3733–3736; *Angew. Chem. Int. Ed.* **2002**, *41*, 3583–3586; c) L. L. Kiessling, C. W. Cairo, *Nat. Biotechnol.* **2002**, *20*, 234–235; d) D. M. Ratner, W. W. Adams, J. Su, B. R. O'Keefe, M. Mrksich, P. H. Seeberger, *ChemBioChem* **2004**, *5*, 379–383.
[22] D. Wang, S. Liu, B. J. Trummer, C. Deng, A. Wang, *Nat. Biotechnol.* **2002**, *20*, 275–281.
[23] B. T. Houseman, M. Mrksich, *Chem. Biol.* **2002**, *9*, 443–454.
[24] E. W. Adams, J. Ueberfeld, D. M. Ratner, B. R. O'Keefe, D. R. Walt, P. H. Seeberger, *Angew. Chem.* **2003**, *115*, 5475–5478; *Angew. Chem. Int. Ed.* **2003**, *42*, 5317–5320.
[25] B. T. Houseman, E. S. Gawalt, M. Mrksich, *Langmuir* **2003**, *19*, 1522–1531.

[26] World Health Organization World Malaria situation 1990, *World Health Stat. Q.* **1992**, *45*, 257–266.
[27] M. C. Hewitt, D. A. Snyder, P. H. Seeberger, *J. Am. Chem. Soc.* **2002**, *124*, 13434–13436.
[28] E. W. Adams, D. M. Ratner, H. R. Bokesch, J. B. McMahon, B. R. O'Keefe, P. H. Seeberger, *Chem. Biol.* **2004**, *11*, 875–881.
[29] C. Walsh, *Nature* **2000**, *406*, 775–781.
[30] G. D. Wright, *Curr. Opin. Chem. Biol.* **2003**, *7*, 563–569.
[31] B. Llano-Sotelo, E. F. Azucena, Jr., L. P. Kotra, S. Mobashery, C. S. Chow, *Chem. Biol.* **2002**, *9*, 455–463.
[32] M. D. Disney, S. Magnet, J. S. Blanchard, P. H. Seeberger, *Angew. Chem.* **2004**, *116*, 1618–1620; *Angew. Chem. Int. Ed.* **2004**, *43*, 1591–1594.
[33] M. D. Disney, P. H. Seeberger, *Chem. Eur. J.* **2004**, *10*, 3308–3314.
[34] S. S. Hegde, F. Javid-Majd, J. S. Blanchard, *J. Biol. Chem.* **2001**, *276*, 45876–45881.
[35] S. Magnet, T. Lambert, P. Courvalin, J. S. Blanchard, *Biochemistry* **2001**, *40*, 3700–3709.
[36] S. S. Hedge, T. K. Dam, C. F. Brewer, J. S. Blanchard, *Biochemistry* **2002**, *41*, 7519–7527.
[37] a) T. J. Baker, N. W. Luedtke, Y. Tor, M. Goodman, *J. Org. Chem.* **2000**, *65*, 9054–9058; b) N. W. Luedtke, T. J. Baker, M. Goodman, Y. Tor, *J. Am. Chem. Soc.* **2000**, *122*, 12035–12036.
[38] M. W. Vetting, S. S. Hegde, F. Javid-Majd, J. S. Blanchard, S. L. Roderick, *Nat. Struct. Biol.* **2002**, *9*, 653–658.
[39] I. Capila, R. J. Linhardt, *Angew. Chem.* **2002**, *114*, 426–450; *Angew. Chem. Int. Ed.* **2002**, *41*, 390–412.
[40] a) B. Casu, *Adv. Carbohydr. Chem. Biochem.* **1985**, *43*, 51–134; b) W. D. Comper, *Heparin and Related Polysaccharides*, *Vol. 7*, Gordon and Breach, New York, **1981**.
[41] a) C. Tabeur, J.-M. Mallet, F. Bono, J.-M. Herbert, M. Petitou, P. Sinaÿ, *Bioorg. Med. Chem.* **1999**, *7*, 2003–2012; b) J.-L. de Paz, J. Angulo, J.-M. Lassaletta, P. M. Nieto, M. Ridondo-Horcajo, R. M. Lozano, G. Gimenez-Gallego, M. Martin-Lomas, *ChemBioChem* **2001**, *2*, 673–685; c) J.-C. Lee, X.-A. Lu, S. S. Kulkarni, Y.-S. Weu, S.-C. Hung, *J. Am. Chem. Soc.* **2004**, *126*, 476–477; d) A. Prabhu, A. Venot, G.-J. Boons, *Org. Lett.* **2003**, *5*, 4975–4978.
[42] H. A. Orgueira, A. Bartolozzi, P. Schell, R. E. J. N. Litjens, E. R. Palmacci, P. H. Seeberger, *Chem. Eur. J.* **2003**, *9*, 140–169.
[43] D. Shukla, J. Liu, P. Blaiklock, N. W. Shworak, X. M. Bai, J. D. Esko, G. H. Cohen, R. J. Eisenberg, R. D. Rosenberg, P. G. Spear, *Cell* **1999**, *99*, 13–22.
[44] R. C. Willis, *Modern Drug Discovery* **2004**, *7*, 36–38.
[45] a) K. A. Karlsson, *Biochem. Soc. Trans.* **1999**, *27*, 471–474; b) K. A. Karlsson, *Adv. Exp. Med. Biol.* **2001**, *491*, 431–443.
[46] M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem.* **1998**, *110*, 2908–2953; *Angew. Chem. Int. Ed.* **1998**, *37*, 2745–2794.
[47] M. D. Disney, J. Zheng, T. M. Swager, P. H. Seeberger, *J. Am. Chem. Soc.* **2004**, *126*, 13343–13346.
[48] M. D. Disney, P. H. Seeberger, *Chem. Biol.* **2004**, *11*, 1701–1707.

Published online: March 30, 2005