Microarrays

DOI: 10.1002/ange.200704998

On-Chip Synthesis and Label-Free Assays of Oligosaccharide Arrays**

*Lan Ban and Milan Mrksich**

Oligosaccharides, like proteins and DNA, are ubiquitous biopolymers that mediate essential functions in organisms.[1] Yet, an understanding of the many roles that carbohydrates play is still at an early stage and essentially absent when compared to our knowledge of the functions of proteins and nucleic acids.[2] This contrast reflects the lack of convenient and flexible tools for the synthesis and biochemical analysis of oligosaccharides and their conjugates. The development of biochips—glass slides patterned with an array of hundreds to thousands of biomolecules—has been particularly important to studies in genomics and proteomics, and first became practical with the on-chip synthesis of oligonucleotides and peptides.[3,4] Herein we present a strategy for the on-chip synthesis of oligosaccharide arrays and we demonstrate that combining these arrays with mass spectrometry permits us to perform label-free assays of glycosyltransferase activities. This method represents a significant addition to the molecular strategies now used to discover and investigate the biological functions of oligosaccharides.

Since our report of a carbohydrate array five years ago,[5] several examples demonstrating the preparation and application of oligosaccaharide arrays have been reported.[6] Wong, Paulson, and co-workers prepared 200 amino-substituted oligosaccharides and immobilized these molecules on *N*-hydroxysuccinimidate-functionalized glass slides for studies of the binding of lectins and viruses.[7] Gildersleeve and co-workers prepared arrays by printing conjugates of carbohydrates and bovine serum albumin and glycoproteins onto epoxide-functionalized glass slides and used these arrays to profile multiple lectins.[8] Seeberger and co-workers used solid phase methods to prepare seven sulfhydryl-terminated oligosaccharides and immobilized these reagents onto maleimide-functionalized glass slides for assaying the interactions between proteins and carbohydrate epitopes.[9] In collaboration with the group of Seeberger we prepared thiol-modified high mannose oligosaccaharides, which were immobilized on self-assembled monolayers, for assays involving lectin binding.[10] The effort required to synthesize the carbohydrate reagents in this and related work is substantial, and limits the size of arrays. Additionally, the need for labeling strategies can make it difficult to perform assays and, further, is not well-suited to the identification of unanticipated activities. Our approach takes advantage of the combination of monolayers with matrix assisted laser desorption-ionization mass spectrometry (in a technique referred to as self-assembled monolayers for matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (SAMDI-TOF MS)) to directly observe the synthetic intermediates at the surface and therefore to optimize the reactions used in the synthesis of the oligosaccharides.[11,12] This method is also useful in performing assays of enzymes that modify the immobilized oligosaccharides and has been applied to nanoparticles.[13]

Our strategy for synthesizing oligosaccharides directly on the biochip is shown in Figure 1 and Table 1, and illustrated in Figure 2 for the synthesis of β-D-galactopyranosyl-1,4-β-D-glucopyranoside (β-D-Gal-1,4-β-D-Glc). The synthesis starts with self-assembled monolayers of alkanethiolates on gold that display a phenol group at a density of 10% amongst the methoxy-terminated tri(ethylene glycol) groups. The phenol group serves as the nucleophile for the attachment of the first carbohydrate building block and the tri(ethylene glycol) groups are effective for preventing the non-specific adsorption of proteins that lead to background signals in subsequent assays. The mass spectrum of this monolayer reveals peaks at

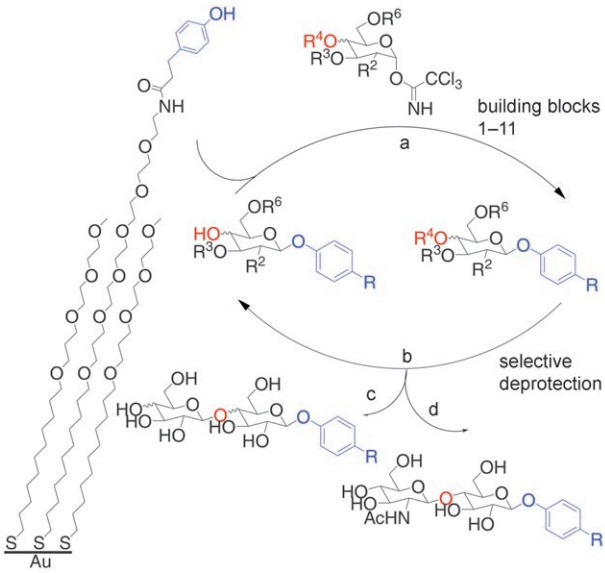

Figure 1. Strategy for the on-chip synthesis of oligosaccharides. The synthesis uses building blocks **1**—**11** that provide for the regioselective coupling of the carbohydrates. a) Attachment of carbohydrate to support. b) Removal of protecting group in preparation for addition of second carbohydrate (repeat step (a)). c) Removal of protecting group(s) to reveal desired disaccharide. d) Removal of protecting group(s) to reveal desired glycosaminosides. R = solid support. Details are explained in the text and the R groups are defined in Table 1.

[*] L. Ban, Prof. M. Mrksich
Howard Hughes Medical Institute and Department of Chemistry
The University of Chicago
Chicago, IL 60637 (USA)
Fax: (+1) 773-702-1677
E-mail: mmrksich@uchicago.edu

[**] The work was supported by the NSF.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

InterScience © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 1: Carbohydrate building blocks.[a]

	R^2	R^3	R^4	R^6
1, **2**	OLev	Ac	Ac	Ac
3, **4**	OAc	Lev	Ac	Ac
5, **6**	OAc	Ac	Lev	Ac
7, **8**	OAc	Ac	Ac	Lev
9, **10**	OAc	Ac	Ac	Ac
11	NHFmoc	Ac	Ac	Ac

[a] **1**, **3**, **5**, **7**, **9**: refer to galactosides; **2**, **4**, **6**, **8**, **10** refer to glucosides; **11** refers to the glucosaminoside.

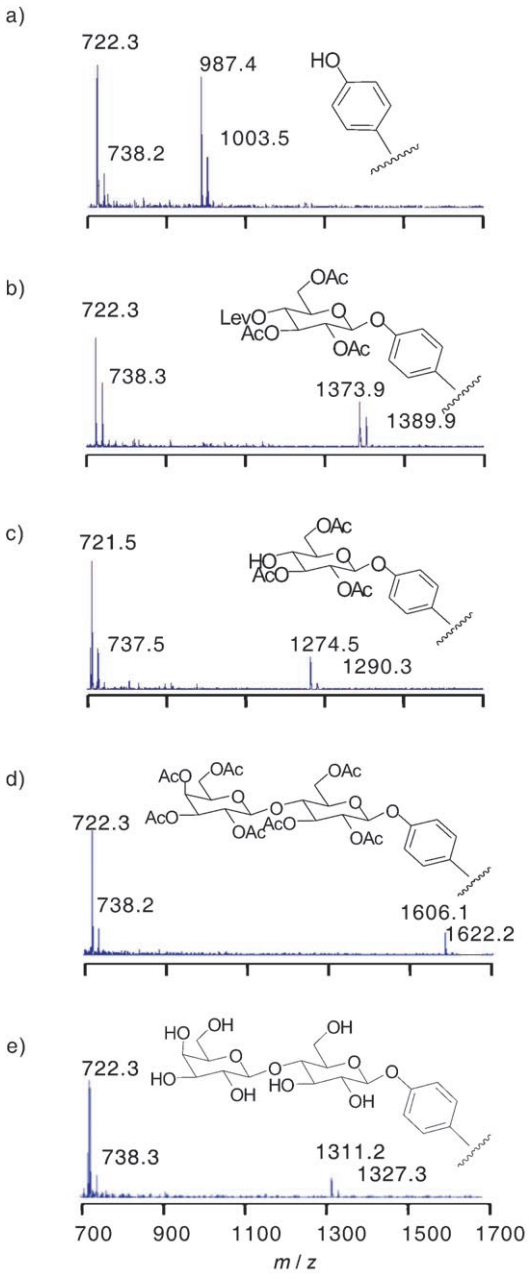

Figure 2. SAMDI-TOF MS results for each intermediate in the synthesis of a disaccharide: a) monolayer displaying the phenol group; b) coupling of the first carbohydrate (**6**); c) selective removal of the levulinate protecting group; d) coupling of the second carbohydrate (**9**); e) final removal of the protecting groups.

987.4 ($C_{50}H_{93}NO_{12}S_2Na$) and 1003.5 ($C_{50}H_{93}NO_{12}S_2K$) for the sodium and potassium adducts, respectively, of the mixed disulfide containing one phenol reactive group and one background tri(ethylene glycol) group. The peaks at 722.3 ($C_{36}H_{74}O_8S_2Na$) and 738.2 ($C_{36}H_{74}O_8S_2K$) correspond to the symmetric disulfide of the background molecules. We prepared 11 monosaccharide building blocks (Table 1) derived from glucose (Glc), galactose (Gal), and *N*-acetyl glucosamine (GlcNAc). The carbohydrates were activated as their trichloroacetimidates[14] and had one of the hydroxy groups protected with the levulinate (Lev)[15] group, and the remaining hydroxy groups were protected with acetyl (Ac) groups. In this way, the carbohydrate could be coupled to the monolayer in the presence of a catalytic amount of trimethylsilyl trifluoromethanesulfonate (TMSOTf) to give the beta anomeric linkage.[12] The monolayer was treated with 2,3,6-tri-*O*-acetyl-4-*O*-levulinyl-α-D-glucopyranosyl trichloroacetimidate (**6**) and TMSOTf to attach the first carbohydrate (Figure 2b). The mass spectrum of the resulting monolayer showed peaks at 1373.9 ($C_{67}H_{115}NO_{22}S_2Na$) and 1389.9 ($C_{67}H_{115}NO_{22}S_2K$) that correspond to the mixed disulfide containing the newly coupled monosaccharide residue and one background tri(ethylene glycol)-terminated alkanethiol. Importantly, the phenol nucleophile reacted completely, as shown by the lack of peaks at *m*/*z* of 987.4 and 1003.5. We then treated the monolayer with hydrazine to remove the Lev group to reveal a site for the attachment of the next carbohydrate building block. The mass spectrum revealed peaks at *m*/*z* 1274.5 ($C_{62}H_{109}NO_{20}S_2Na$) and 1290.3 ($C_{62}H_{109}NO_{20}S_2K$) that correspond to the anticipated intermediate (Figure 2c). Treatment of the monolayer with the second building block, 2,3,4,6-tetra-*O*-acetyl-α-D-galactopyranosyl trichloroacetimidate (**9**), resulted in the protected disaccharide as evidenced by peaks at *m*/*z* 1606.1 ($C_{76}H_{127}NO_{29}S_2Na$) and 1622.2 ($C_{76}H_{127}NO_{29}S_2K$) (Figure 2d). Finally, the monolayer was treated with sodium methoxide (NaOMe) to remove the acetoxy groups and to produce the free disaccharide. The mass spectrum for this final step shows peaks at *m*/*z* 1311.2 ($C_{62}H_{113}NO_{22}S_2Na$) and 1327.3 ($C_{62}H_{113}NO_{22}S_2K$) (Figure 2e). For oligosaccharides that contain glycosaminosides, we used building block **11**, in which the primary amine was protected with the 9-fluorenylmethoxycarbonyl (Fmoc) group, because activation of trichloroacetimide donors having a 2-*N*-acyl group generates relatively unreactive oxazoline intermediates.[16] For these cases, the acetamide group was installed after coupling of the sugar by first removing the Fmoc group with piperidine and then acylating the resulting amine with acetic anhydride (Ac_2O). The mass spectrum representing one such case is shown in Figure 3a.

To prepare the array we started with a monolayer that was patterned with a 4×6 grid of circles that each displayed the phenol group at a 10% density amongst the methoxy-terminated tri(ethylene glycol) groups and the intervening regions displayed a methyl-terminated monolayer (see the Supporting Information for further descriptions). The hydrophobic surface surrounding the circles prevents the spread of reagents that are applied to the circular regions and therefore confines the chemical reactions to each region. We prepared

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

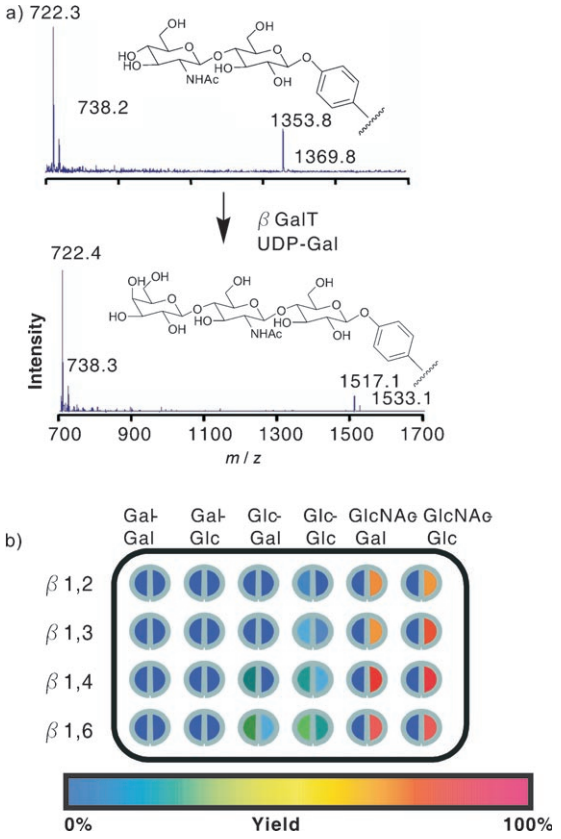

Figure 3. β-GalT activity assays on the 24 disaccharide array. a) SAMDI-TOF MS spectrum for starting material β-D-GlcNAc-1,4-β-D-Glc and product β-D-Gal-1,4-β-D-GlcNAc-1,4-β-D-Glc. b) Activities of β-GalT for each disaccharide. For each circle, the two halves represent results from two different reaction conditions for the same disaccharide. The left half circle represents activity in the presence of added lactalbumin to the enzymatic solution and the right half circle represents the activity in the absence of added lactalbumin.

an array of 24 disaccharides by using either Gal or Glc as the first residue and Gal, Glc, or GlcNAc as the second residue. By using four distinct building blocks each for Gal and Glc in the first step—wherein the building blocks differed in the position of the Lev group—and one building block each for Gal, Glc, and GlcNAc in the second step, the array presented the six disaccharides Gal–Gal, Gal–Glc, Glc–Gal, Glc–Glc, GlcNAc–Gal, and GlcNAc–Glc having each of the four linkages (β 1,2, β 1,3, β 1,4, and β 1,6). The reaction time for the coupling of the carbohydrate building blocks is approximately 30 minutes and the time required to remove the Lev group is one minute. The final treatment to remove the protecting groups and reveal the free disaccharides required one hour; therefore the total time required to prepare the array was about four hours. The mass-spectrometry results showed that each disaccharide was formed as expected and no peaks for the starting material remained, implying high yields for the conversions.

We used the array to profile the substrate specificity of bovine β-1,4-galactosyltransferase I (β-GalT). This enzyme transfers galactoside from uridine diphosphogalactose (UDP-Gal) to its substrate; it is involved in lactose synthesis and in the biosynthesis of glycolipids and glycoproteins.[17] β-GalT has the interesting characteristic that it recognizes *N*-acetyl glucosamine as its substrate, but in the presence of lactalbumin it has an altered specificity and prefers glucose as the substrate.[18] Our array presents three different terminal carbohydrates (Gal, Glc, and GlcNAc) in the context of a panel of disaccharides and is therefore well-suited for assessing the extent to which the identity of the distal sugar and the regiochemical linkage influences the activity of the substrate. We treated disaccharides in the array with a reaction cocktail (5 μL) that included β-GalT (20 U mL^{-1}), $MnCl_2$ (10 mM), HEPES buffer (20 mM, pH 7.6), UDP-Gal (0.5 mM) and in certain experiments, lactalbumin (0.2 mg mL^{-1}). The reactions were run for two hours in a humidified chamber at 37 °C, stopped by rinsing the monolayers with water, and analyzed by SAMDI-TOF MS techniques to give a spectrum for each spot in the array.

Prior to treatment of the array with the enzyme, the spot that represents β-D-GlcNAc-1,4-β-D-Glc showed the expected peaks at *m*/*z* 1353.6 and 1369.8 that correspond to the sodium and potassium adducts (Figure 3a), respectively, of the disulfide containing the disaccharide-terminated alkanethiolate. After treatment with β-GalT, SAMDI-TOF MS revealed new peaks at *m*/*z* of 1517.1 and 1533.1, which correspond to the expected trisaccharide product. The absence of the original peaks at *m*/*z* of 1353.6 and 1369.8 demonstrated that the enzymatic reaction was essentially complete. We could estimate the approximate reaction yield by taking the ratio of the sum of the intensities of product peaks (the trisaccharide) and the sum of the intensities of the product peaks and the reactant peaks (the disaccharide). We performed this reaction on the full array and determined reaction yields for each of the 24 disaccharides in the presence and absence of lactalbumin (Figure 3b). We recognize that these yields represent relative rather than absolute measures since the products may have different ionization efficiencies relative to the substrates.

We found that each of the eight GlcNAc-terminated dissacharides was converted into the respective trisaccharide in high yield. The addition of lactalbumin inhibited the galactosylation of GlcNAc-terminated disaccharides, and instead promoted addition of UDP-Gal to Glc-terminated disaccharides. In the absence of lactalbumin, these same Glc-terminated substrates were mostly inactive. We found that the yields for galactosylation of the Glc-terminated disaccharides (in the presence of lactalbumin) were substantially lower than the yields for GlcNAc-terminated substrates (without lactalbumin), in agreement with previous reports.[17,18] A comparison of the relative activities of all disaccharides revealed that those having β-1,6 and β-1,4 linkages were more active, likely owing to a better fit of the substrate within the enzyme active site.[19,20] Finally, the Gal-terminated disaccharides showed no activity, either with or without lactalbumin, which confirms previous results that the enzyme does not use terminal galactosides as substrates.[18]

This work addresses the significant need for tools that can facilitate the discovery and understanding of the many roles that carbohydrates play in biology. At the molecular level, these needs include the elucidation of binding preferences of

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

proteins that recognize carbohydrates, the substrate specificities of enzymes that are involved in biosynthesis or regulation of carbohydrate epitopes, and the discovery of novel carbohydrate-directed activities of proteins. These studies require appropriately designed libraries of oligosaccharides and convenient assays for assessing relative activities. Significant progress in solid-phase methods for the synthesis of complex oligosaccharides has provided access to a broad range of carbohydrates, and has been critical to recent advances in glycobiology.[21] Yet, the expense and time associated with preparing and assaying oligosaccharides still limits the routine use and large-scale use of biochip arrays. Our report of a method to prepare arrays of oligosaccharides by directly assembling the carbohydrates onto the biochip combined with label-free assay formats enables a rapid and efficient strategy to investigate the biological roles of carbohydrates and their associated proteins. We believe that the combination of on-chip synthesis and label-free assays demonstrated herein will have a significant impact in the study of carbohydrate function. However, this technique is still at an early stage of development and has limitations; for example, the inability to directly determine the configuration of the anomeric linkage in carbohydrates on the monolayer. We expect that additional work will address this theme and will bring increased throughput to the synthesis and application of arrays, ultimately bringing to glycomics the same tools that have transformed the genomic and proteomic sciences.

Experimental Section

Complete schemes and details for the synthesis of carbohydrate building blocks **1**–**11** and phenol terminated disulfide **38**, carbohydrate array fabrication, direct on-chip SAMDI-TOF MS methods, and β-1,4-galactosyltransferase activity assay and quantification are given in the supporting information.

Received: October 29, 2007
Revised: January 10, 2008
Published online: March 26, 2008

Keywords: biochips · enzymes · glycosides · mass spectrometry

[1] C. R. Bertozzi, L. L. Kiessling, *Science* **2001**, *291*, 2357 – 2364.
[2] P. H. Seeberger, D. B. Werz, *Nature* **2007**, *446*, 1046 – 1051.
[3] S. P. Fodor, J. L. Read, M. C. Pirrung, L. Stryer, A. T. Lu, D. Solas, *Science* **1991**, *251*, 767 – 773.
[4] R. Frank, *Tetrahedron* **1992**, *48*, 9217 – 9232.
[5] B. T. Houseman, M. Mrksich, *Chem. Biol.* **2002**, *9*, 443 – 454.
[6] A. S. Culf, M. Cuperlovic-Culf, R. J. Quellett, *OMICS* **2006**, *10*, 289 – 310.
[7] O. Blixt, et al., *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17033 – 17038, see the Supporting Information.
[8] J. C. Manimala, T. A. Roach, Z. Li, J. C. Gildersleeve, *Angew. Chem.* **2006**, *118*, 3689 – 3692; *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 3607 – 3610.
[9] 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.
[10] D. M. Ratner, E. W. Adams, J. Su, B. R. O'Keefe, M. Mrksich, P. H. Seeberger, *ChemBioChem* **2004**, *5*, 379 – 383.
[11] J. Su, M. Mrksich, *Langmuir* **2003**, *19*, 4867 – 4870.
[12] J. Li, P. S. Thiara, M. Mrksich, *Langmuir* **2007**, *23*, 7223 – 7231.
[13] H. Shimizu, M. Sakamoto, N. Nagahori, S.-I. Nishimura, *Tetrahedron* **2007**, *63*, 2418 – 2425.
[14] R. R. Schmidt, J. Michel, *Angew. Chem.* **1980**, *92*, 763 – 764; *Angew. Chem. Int. Ed. Engl.* **1980**, *19*, 731 – 732.
[15] J. H. Van Boom, P. M. Burgers, *Tetrahedron Lett.* **1976**, *17*, 4875.
[16] V. K. Srivastava, *Carbhydr. Res.* **1982**, 103, 286 – 292.
[17] B. Ramakrishnan, P. S. Shah, P. K. Qasba, *J. Biol. Chem.* **2001**, *276*, 37665 – 37671.
[18] F. L. Schanbacher, K. E. Ebner, *J. Biol. Chem.* **1970**, *245*, 5057 – 5061.
[19] L. N. Gastinel, C. Cambillau, Y. Bourne, *EMBO J.* **1999**, *18*, 3546 – 3557.
[20] I. Brockhausen, M. Benn, S. Bhat, S. Marone, J. G. Riley, P. Montoya-Peleaz, J. Z. Vlahakis, H. Paulsen, J. S. Schutzbach, W. A. Szarek, *Glycoconjugate J.* **2006**, *23*, 525 – 541.
[21] For a review, see: L. A. Marcaurelle, P. H. Seeberger, *Curr. Opin. Chem. Biol.* **2002**, *6*, 289 – 296.

© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim