Carbohydrate Microarrays

Carmen Ortiz Mellet[b] and José M. García Fernández*[a]

KEYWORDS:

antibodies · antigens · carbohydrates · lectins · microarrays

Glycan transfer to protein chains and the subsequent trimming and processing of the attached oligosaccharides represents a critically important posttranslational modification reaction and is a target for future proteomic research. When conjugated to a protein to form glycoproteins, oligosaccharides can alter protein folding and charge (for example, by sialylation or sulfation) and induce heterogeneous profiles as a consequence of differing glycoforms, thus providing a mechanism to modulate the protein's behavior in a complex multicellular environment. Glycosylated proteins are ubiquitous components of extracellular matrices and cellular surfaces, where their oligosaccharide moieties are involved in extensive recognition phenomena including development, differentiation, morphogenesis, fertilization, the immune response, implantation, cell migration, and cancer metastasis.[1] Every cell in every living organism is covered with an abundance of such diverse carbohydrate chains, the glycocalyx, whose composition reflects not only cell types but also different cell states. Interestingly, abnormalities in protein glycosylation have often been correlated with specific disease states. This has lead to the development of therapeutic agents designed to interfere with carbohydrate biosynthesis or molecular recognition[2] and carbohydrate-based anticancer vaccines.[3] The complementarity of defined sugar epitopes and specific protein receptors has also found applications in vectorized drug or probe delivery[4] and in the targeted aggregation of pathogenic species.[5]

Progress in this field, for which the term *glycomics* has been coined, is limited by the extremely broad diversity of carbohydrate structures and the different contexts in which they occur. Although this structural diversity renders carbohydrates ideally suited to the transfer of biological information, it also it makes difficult to dissect the individual roles of each receptor – ligand pair. Automated chip-based technologies for rapidly and quantitatively assessing interactions between large numbers of oligosaccharide structures and proteins simultaneously, analogous to the complementary DNA (cDNA) chip-based technologies that have facilitated transcriptomics,[6] could provide a leap forward in glycomics research.[7] The work by Wang et al.[8] at Columbia University, by Park and Shin[9] at Yonsei University, Seoul, and by Houseman and Mrksich[10] at The University of Chicago represents a decisive step in that direction. The first group describe the fabrication of a carbohydrate-based microarray and demonstrate that it can be successfully used to identify carbohydrate antigenic determinants and to detect the presence of specific complementary antibodies in a given sample, including human serum. The last two groups have developed mono- and disaccharide chips as suitable tools for studying high-throughput carbohydrate – lectin interactions.

Glycoproteins play an active role in the innate and adaptive immune response and, in many cases, specific glycoforms are involved in the immune process.[11] Cells of the immune system use the glycans on the surface of the cells that they encounter to identify everything from bacteria to partners. Changes in glycoprotein glycosylation pattern may lead, for instance, to inappropriate activation of the innate immune system, which is in the origin of autoimmune diseases such as rheumatoid arthritis. Similarly, on tumor cells aberrant glycosylation can expose new saccharide epitopes for recognition by the immune system. The diverse repertoire of oligosaccharides displayed on the surface of a cell can be exploited by viruses that first adhere and then use the host glycosylation machinery to help avoid immune detection. The benefits of understanding protein – carbohydrate interactions and identifying carbohydrate epitopes that are important for specific recognition events are, thus, evident. The problem has been addressed by intense multidisciplinary research in glycobiology. Despite important advances, the rate of generating information, in comparison with genomics or proteomics research, is slow due to the cumbersome multistep methods currently available and the lack of high-throughput methods for characterizing protein – carbohydrate interactions.

To be useful for biomedical research and clinical diagnosis, any method for detecting and characterizing carbohydrate recognition must achieve the sensitivity to detect binding to specific receptors when the amount of specimen is limited, thereby circumventing the problem of intrinsic weak affinities between carbohydrates and their target proteins.[12] The microarray platform looks well adapted to address this issue. Presentation of sugars on surfaces and monolayers creates a multivalent display that efficiently mimics the natural mode of affinity enhancement that arises from multiple interactions between the binding proteins and the carbohydrate ligands. Several methodologies for the immobilization of carbohydrates and glycoconjugates on surfaces for binding studies have been

[a] *Dr. J. M. García Fernández*
Instituto de Investigaciones Químicas, CSIC
Américo Vespucio s/n
Isla de la Cartuja
41092 Sevilla (Spain)
Fax: (+34) 9544-60565
E-mail: jogarcia@cica.es

[b] *Dr. C. Ortiz Mellet*
Departamento de Química Orgánica
Facultad de Química
Universidad de Sevilla
Apartado 553, 41071 Sevilla (Spain)

  1439-4227/02/03/09 $ 20.00+.50/0

reported.[13] Probably the simplest one is the known adsorption of some polysaccharides and carbohydrate polymers onto the plastic surface of a microtiter well through noncovalent interactions, which is the basis of the enzyme-linked lectin assay (ELLA) protocol.[14] Wang et al.[8] have adopted a similar strategy for the fabrication of glyco-arrays: glycans (48 compounds derived from biological sources) were spotted onto glass slides precoated with nitrocellulose polymer to which they attach, without the need for chemical conjugation. The air-dried printed carbohydrate microarrays can be stably stored at room temperature and can be used after just blocking with bovine serum albumin (BSA) as an irrelevant protein.

Two prior conditions for the feasibility of the glyco-array approach in assessing carbohydrate – protein interactions are: (i) a high efficiency of the immobilization technique and (ii) the immobilized carbohydrate-containing macromolecules must preserve their recognition properties. The second point is not trivial; as shown by the groups of Kahne and Whitesides,[15] secondary interactions in carbohydrate-derivatized monolayers may lead to a switch in carbohydrate-binding selectivity. To address these questions, Wang et al.[8] produced a microarray by printing a series of fluorescein isothiocyanate (FITC) conjugated dextrans differing in molecular mass and structure on the modified glass slides. By analyzing the fluorescence intensities retained after extensive washing, they concluded that polysaccharides of 3.3 – 2000 kDa were all stably immobilized. Yet the efficiency of their immobilization was significantly influenced by the molecular weight, with the bigger molecules being better retained. Most importantly, the printed dextran preparations were specifically recognized by monoclonal antibodies of defined specificities and exhibited antigen – antibody reactivity patterns identical to those obtained by classical quantitative immunoassays such as ELISA.

These results paved the way to further explore the high-throughput nature of the carbohydrate microarray technology. By using a printed panel of 48 carbohydrate-containing macromolecules, several distinct specificities of anti-carbohydrate antibodies present in the human serum of normal individuals were identified (Figure 1), including antibodies that bound to pathogens such as *Escherichia coli*, *Pneumococcus*, *Meningococcus*, or *Haemophilus influenzae*.

In this particular study, most of the detected glycan – antibody positive bindings were consistent with the expected specificity. What makes the work by Wang et al.[8] still more exciting is that they have demonstrated that the carbohydrate microarray construction can also be used to discover and characterize unexpected biologically relevant antibody specificities. Thus, the monoclonal antibody 4.3F1, which recognizes terminal (nonreducing) α-(1$\rightarrow$6)-linked D-glucopyranose subunits, was observed to interact with chondroitin sulfate B derived from pig intestinal mucosa and lacking this epitope. To discard the possibility that this hit was an artefact of the in vitro microarray assay, further experiments with intestinal tissue from pigs and mice were performed. The antiterminal α-(1$\rightarrow$6)-glucose antibody 4.3F1 clearly stained a subpopulation of monocyte/macrophages in the lamina propria of the small intestine, therefore validating the result obtained with the microarray platform.

The work of Wang et al.[8] focused on the use of carbohydrate microarrays for detection of carbohydrate antigen – antibody interactions. A direct application of the method would be the rapid diagnosis of infectious diseases. Other potential uses include detection of tumor cells, identification and profiling of novel carbohydrate-binding proteins or carbohydrate-processing enzymes, and identification of novel inhibitors of glycan – protein interactions. More robust immobilization techniques may be necessary for those channels, especially when low-molecular-weight recognition epitopes are involved. The approach disclosed by Park and Shin[9] to fabricate mono- and disaccharide microarrays comes to fill this gap by covalently attaching the carbohydrate ligands to thiol-derivatized glass slides. The proposed three-step reaction sequence involves: (i) formation of a glycosylamine, (ii) subsequent coupling with an

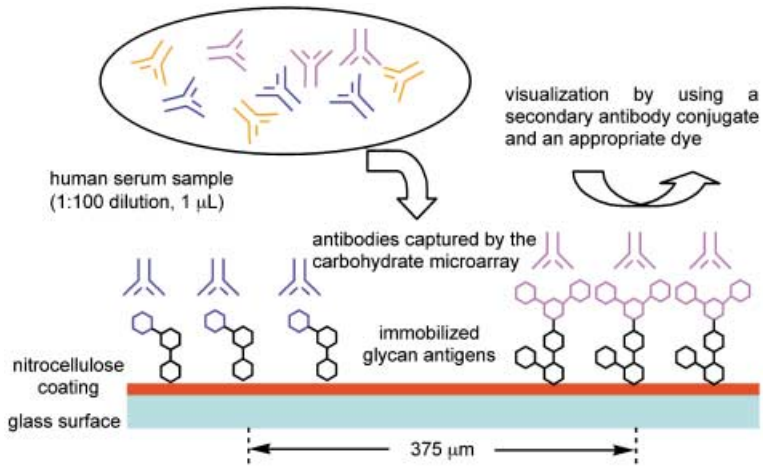
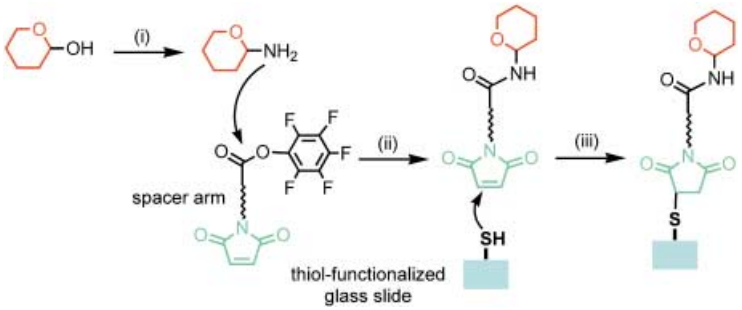


Figure 1. *Schematic representation of the microarray device conceived by Wang et al.[8] Carbohydrate antigens were spotted onto the glass slides precoated with nitrocellulose polymer with the help of a high-precision robot designed to produce cDNA microarrays. The immobilized carbohydrate-containing macromolecules include:* Klebsiella *capsular polysaccharides (types K7, K11, K12, K13, K14, K33, and A3), Dudman's* Rhizobium tritolii *A1 polysaccharide, chondroitin sulfate (A, B, and C),* Pneumococcal *C polysaccharide,* Pneumococcus *capsular polysaccharide (types SIV, VIII, IX, XIV, and 27), cow blood group B (cow 21 and cow 26), Bacto-agar (20 °C extract), arabino galactan (Larch CORASH), isomaltotriose (BSA and keyhole limpet hemocyanin (KLH) conjugates), Lewis^a blood group substance, human ovarian cyst blood group substance (Beach P1, Tij II, and group A), blood group substance I (Ogunsheye 10% 2X), asialo-orosomucoid, lacto-*N*-tetraose (BSA conjugate), Phosphomannan NRRL B-2448,* Menningococcus *capsular polysaccharide (groups B and Y),* H. influenzae *capsular polysaccharide (type A),* E. coli *capsular polysaccharide (types K1, K92, and K100), D-galactan (from* H. pomiata *and* H. nemoralis*), isomaltohexaose (BSA and KLH conjugates), dextran N-150-N (average* M_w *60 kDa), Hog blood group O substance (H), agalacto orosomucoid, inulin, and levan (from B-512-E dextran). Postive hits were visualized by using either anti-human IgM-AP conjugate and Vector Red or biotinilated anti-human IgG and Cy3-streptavidin. The stained microarrays were scanned with a standard biochip scanning system.*

activated ω-(*N*-maleoyl)carboxylic acid derivative, and (iii) Michael addition of the thiol group on the solid support to the maleimide fragment in the armed ligand. Finally, the unreacted thiol groups are capped by reaction with excess *N*-ethylmaleimide (Scheme 1).

The incorporation of a long-enough spacer (C_6 or longer) between the covalently attached sugar and the maleimide thioether anchor seems to be an important requisite to warrant accessibility of these relatively small molecules to recognition events, a phenomenon already observed in highly dense glycoclusters.[16] In any case, probing of microspotted slides containing appropriately tethered *N*-acetyl-D-glucosamine, lactose, maltose, and cellobiose ligands with FITC-tagged lectins showed a binding pattern identical to that in solution (Figure 2); this demonstrates that microarrays prepared by this technique allow thousands of carbohydrate–protein binding assays (up to 12000 microspots in this work) to be performed in a parallel fashion. Moreover, the chip can be recycled without detaching of the carbohydrate ligands.

The examples given above show that carbohydrate microarrays are useful tools to study different kinds of carbohydrate interactions. Yet, different immobilization techniques, such as in situ synthesis of the oligosaccharides[14b] or the generation of self-assembled monolayers (SAMs) onto gold surfaces from sulfur-containing conjugates,[11a, 15] might be necessary for optimization of array production and assay performance. The latter approach has recently been implemented by Houseman and Mrksich[10] to generate monosaccharide glyco-arrays that exhibit a very good control of ligand density and that are compatible with different detection methods, including fluorescence microscopy and surface plasmon resonance. Their methodology to engineer the gold surface for immobilization of the carbohydrate ligands is depicted in Scheme 2. Hydroquinone-containing SAMs are first produced and then oxidized to the corresponding benzoquinone derivatives to which cyclopentadiene conjugates are covalently attached through a Diels–Alder cycloaddition reaction.



Scheme 1. *The key steps for the preparation of carbohydrate microarrays according to Park and Shin[9] methodology: (i) direct formation of a glycosylamine by treatment of the free sugars (*N*-acetyl-D-glucosamine (GlcNAc), lactose, maltose and cellobiose) with aqueous ammonium hydrogencarbonate/ammonium hydroxide; (ii) coupling of the glycosylamine with the pentafluorophenyl ester of an* ω-N-*maleoyl)carboxylic acid derivative (the spacer arms used range from C_2 to C_{24}); and (iii) Michael addition of the thiol groups on the modified glass slide to the double bond of the maleimide fragment in the conjugate. The last reaction was accomplished in 5 h after printing at room temperature. The unreacted thiol groups were than capped by reaction with excess* N-*ethylmaleimide.*

As the Diels–Alder reaction is rapid, selective, and quantitative, all carbohydrates within the array are presented at a uniform density. Further advantages of this system are the possibility of quantitatively determining the density of reactive quinone groups from their reversible electrochemical reduction and the inertness to nonspecific adsorption of proteins, with no need for the BSA-blocking treatment; the latter provides excellent signal-to-noise responses. Nonreacted benzoquinone groups are inactivated by treatment with a tri(ethylene glycol)–cyclopentadiene derivative. A microarray containing ten different monosaccharides was prepared by this technique and assayed against a set of lectins incorporating a fluorescent dye. The observed specificities exquisitely matched those observed in solution.

The above ten-spot glycochip example is far from satisfying the requirements for high-throughput analysis of carbohydrate–protein interactions that research in glycomics demands. Nevertheless, the authors have already foreseen the compatibility of the present design with the automated solid-phase synthesis of complex *n*-pentenyl glycosides according to the method of Seeberger's group.[17] The terminal olefin can be easily transformed

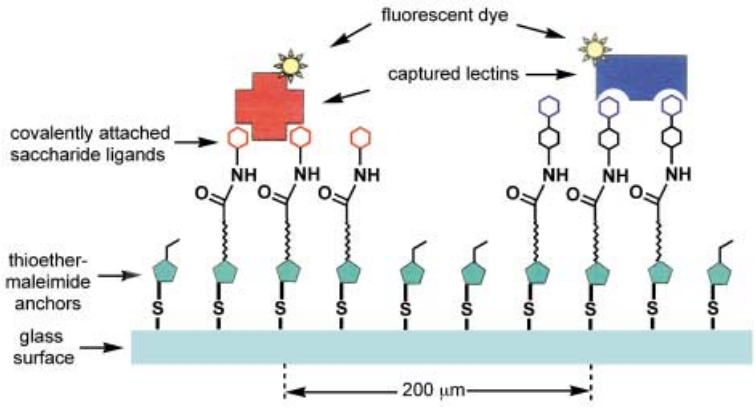
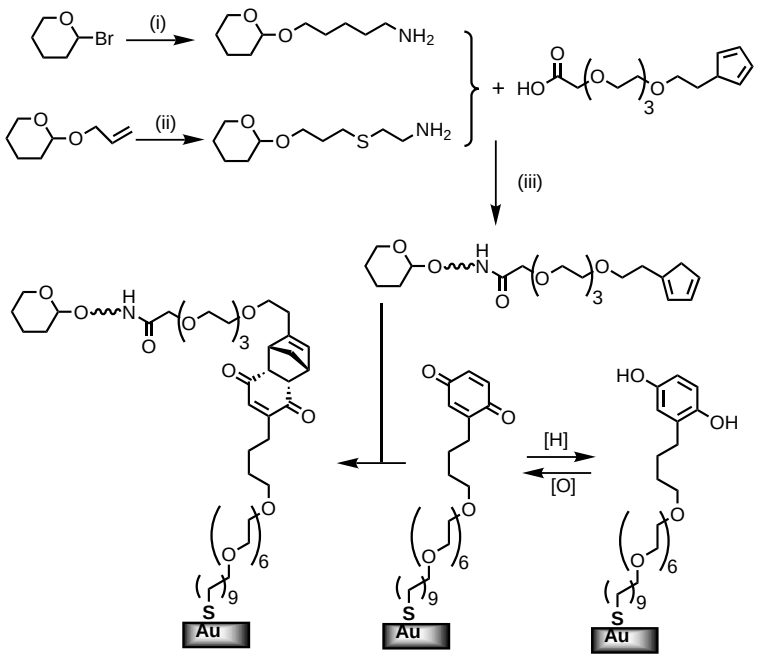


Figure 2. *Schematic representation of the carbohydrate chip for high-throughput detection of carbohydrate–lectin interactions developed by Park and Shin.[9] The slides (7.5 × 2.5 cm, up to 12000 microspots) were probed with fluorescein-labeled lectins (Concanavalin A (ConA),* Erythrina cristagalli *(EC), and* Triticum vulgaris *(TV)) and scanned after removal of the unbound lectins by extensive washing. The recognition pattern in the microchip matched that in solution for the assayed sugars and lectins, that is, TV specifically recognized GlcNAc, EC lactose, and ConA maltose. As expected, cellobiose was not recognized by any of these lectins.*



Scheme 2. *The methodology of Houseman and Mrksich[10] for the preparation of glyco-arrays onto SAMs. Titanium (5 nm) and then gold (15 nm) were evaporated onto glass coverslips which were subsequently immersed in a methanolic solution containing hydroquinone-terminated alkanethiol (10 μM) and penta-(ethylene glycol)-terminated alkanethiol (1 mM total thiol). Appropriately armed monosaccharide ligands (α- and β-glucose, α- and β-galactose, α- and β-fucose, α- and β-GlcNAc, α-mannose, and α-rhamnose) were prepared by: (i) glycosylation of the corresponding peracetylated glycosyl bromide with a* N*-protected amino alcohol by using $Hg(CN)_2$ as a promotor or (ii) photochemical addition of cysteamine to α-allyl glycosides, followed by (iii) acylation of the terminal amino group with a cyclopentadiene-terminated carboxylic acid. To prepare the arrays, the immobilized hydroquinone groups were oxidized to the corresponding benzoquinone by treatment with a saturated aqueous solution of 1,4-benzoquinone. Then, 1 mL of each monosaccharide conjugate (2 mM in H_2O) was applied to specific regions of the monolayer. The coupling reaction was accomplished in 2 h at 37 °C in a humidified chamber.*

to an aldehyde group and then chemoselectively ligated to an hydrazide – tri(ethylene glycol) – cyclopentadiene arm. Interestingly, the immobilized carbohydrates can be further subjected to enzymatic elaboration, as demostrated in a particular example: GlcNAc ligands in the array were transformed into *N*-acetyllactosamine ligands by the action of bovine β-1,4-galactosyltransferase and UDP-galactose. That also illustrates the potential of this platform in the profiling of carbohydrate processing enzymes.

The coalescence of increasingly efficient techniques for the analysis of glycan structure and function, such as ultrahigh-sensitive mass spectrometric methods,[18] with emerging improved means for the isolation[19] or synthesis[20] of carbohydrate libraries, will certainly accelerate the upcoming of much larger glyco-arrays; these in turn will extend the number of applications dramatically. Glycomics research, diagnostic applications, and drug discovery will be the major fields addressed by carbohydrate microarray technology. The story is just beginning. Yet the new generations of carbohydrate microarrays are likely to impact on our molecular understanding of how the entire set of glycans in an organism, the *glycome*, mediates physiological and pathological events.

[1] a) K. M. Koeller, C.-H. Wong, *Nat. Biotechnol.* **2000**, *18*, 835 – 841; b) R. S. Singh, A. K. Tiwary, J. F. Kennedy, *Crit. Rev. Biotechnol.* **1999**, *19*, 145 – 178.

[2] a) C. R. Bertozzi, L. L. Kiessling, *Science* **2001**, *291*, 2357 – 2364; b) P. I. Kitov, J. M. Sadowska, G. Mulvey, G. D. Armstrong, H. Ling, N. S. Pannu, R. J. Read, D. R. Bundle, *Nature* **2000**, *403*, 669 – 672; b) P. Sears, C.-H. Wong, *Angew. Chem.* **1999**, *111*, 2446 – 2471; *Angew. Chem. Int. Ed.* **1999**, *38*, 2300 – 2334.

[3] a) S. J. Danishefsky, J. R. Allen, *Angew. Chem.* **2000**, *112*, 882 – 912; *Angew. Chem. Int. Ed.* **2000**, *39*, 836 – 863; b) S. Hakomori, Y. Zang, *Chem. Biol.* **1997**, *4*, 97 – 104.

[4] C. Ortiz Mellet, J. Defaye, J. M. García Fernández, *Chem. Eur. J.* **2002**, *8*, 1982 – 1990.

[5] J. E. Gestwicki, L. E. Strong, C. W. Cairo, F. J. Boehm, L. L. Kiessling, *Chem. Biol.* **2002**, *9*, 163 – 169.

[6] A. Shulze, J. Downward, *Nat. Cell Biol.* **2001**, *3*, E190-E195.

[7] A. Dove, *Nat. Biotechnol.* **2001**, *19*, 913 – 917.

[8] D. Wang, S. Liu, B. J. Trummer, A. Wang, *Nat. Biotechnol.* **2002**, *20*, 275 – 281.

[9] S. Park, I. Shin, *Angew. Chem.* **2002**, *114*, 3312 – 3314; *Angew. Chem. Int. Ed.* **2002**, *41*, 3180 – 3182.

[10] B. T. Houseman, M. Mrksich, *Chem. Biol.* **2002**, *9*, 443 – 454.

[11] P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson, R. A. Dwek, *Science* **2001**, *291*, 2370 – 2375.

[12] a) J. C. Sacchettini, L. G. Baum, C. F. Brewer, *Biochemistry* **2001**, *40*, 3009 – 3015; b) R. T. Lee, Y. C. Lee, *Glycoconjugate J.* **2000**, *17*, 543 – 551.

[13] a) B. T. Houseman, M. Mrksich, *Top. Curr. Chem.* **2002**, *218*, 1 – 44; b) C. Leteux, R. A. Childs, W. Chai, M. S. Stoll, H. Kogelberg, T. Feizi, *Glycobiology* **1998**, *8*, 227 – 236.

[14] a) J. J. Lundquist, E. Toone, *Chem. Rev.* **2002**, *102*, 555 – 578; b) L. S. Khratsova, M. A. Sablina, T. D. Melikhova, D. H. Joziasse, H. Kaltner, H.-J. Gabius, N. Bovin, *Anal. Biochem.* **2000**, *280*, 250 – 257.

[15] N. Horan, L. Yan, H. Isobe, G. M. Whitesides, D. Kahne, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11 782 – 11 786.

[16] I. Baussanne, J. M. Benito, C. Ortiz Mellet, J. M. García Fernández, J. Defaye, *ChemBioChem* **2001**, *1*, 777 – 783.

[17] O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science*, **2001**, *291*, 1523 – 1527.

[18] a) Y. Mechref, M. V. Novotny, *Chem. Rev.* **2002**, *102*, 321 – 369; b) A. Dell, H. Morris, *Science* **2001**, *291*, 2351 – 2356.

[19] J. Hirabayashi, Y. Arata, K. Kasai, *Proteomics* **2001**, *1*, 295 – 303.

[20] a) P. Sears, C.-H. Wong, *Science* **2001**, *291*, 2344 – 2350; b) P. M. St. Hilaire, M. Meldal, *Angew. Chem.* **2000**, *112*, 1210 – 1228; *Angew. Chem. Int. Ed.* **2000**, *39*, 1162 – 1179.