Development of a Lectin Microarray for the Rapid Analysis of Protein Glycopatterns

Kanoelani T. Pilobello,[a]
Lakshmipriya Krishnamoorthy,[b] Deepika Slawek,[a]
and Lara K. Mahal*[c]

To A.F. in memoriam

Glycosylation plays a role in a wide variety of biological processes including bacterial pathogenesis, tumor cell metastasis, and inflammation.[1] Over half of all proteins are believed to be glycosylated; this makes carbohydrates one of the most common post-translational modification motifs.[2] Despite both the ubiquitous nature of carbohydrates and their importance, the heterogeneity and complexity of protein glycoforms have impeded study of their function.[3] Increased interest in the systematic evaluation of glycosylation has led to the creation of a new field, glycomics.[4–7] Methods in this field thus far have focused on the creation and use of carbohydrate arrays, in which a variety of oligosaccharides, glycolipids, or glycoproteins are bound to solid supports and used to probe the carbohydrate-binding properties of proteins or cells.[6,8–13] Although these carbohydrate arrays yield valuable information about carbohydrate-interacting proteins, they do not allow us to directly examine changes in glycosylation. Alterations in the carbohydrate composition of glycoproteins are known to coincide with changes in protein clearance, cell-adhesion properties, and tumor cell states.[1,14] Current technologies available for glycan analysis, such as mass spectrometry, Western blotting, and chromatography, tend to be time-consuming and ill-suited to the rapid and systematic evaluation of protein glycosylation states. Although mass spectrometry, in particular, gives highly detailed carbohydrate structures, the time and expertise required make this technique difficult for the average researcher to access. Herein, we describe the development of a lectin microarray for the rapid and simple survey of protein glycosylation.

[a] *K. T. Pilobello, D. Slawek*
Department of Chemistry and Biochemistry
University of Texas at Austin
1 University Station, A5300
Austin, TX 78712-0265 (USA)

[b] *L. Krishnamoorthy*
Institute for Cellular and Molecular Biology
University of Texas at Austin
1 University Station, A5300
Austin, TX 78712-0265 (USA)

[c] *Dr. L. K. Mahal*
Department of Chemistry and Biochemistry and
the Institute for Cellular and Molecular Biology
University of Texas at Austin
1 University Station, A5300, Austin, TX 78712-0265 (USA)
Fax: (+1) 512-471-8696
E-mail: lmahal@cm.utexas.edu

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

Microarrays enable a multitude of discrete interactions to be observed simultaneously, thus resulting in a probe-based profile of the sample.[15,16] Lectins are carbohydrate-binding proteins that have been used as probes for glycan detection for decades in a wide variety of biological assays, including histology, flow cytometry, blotting, and, recently, as a means to differentiate bacterial subtypes for biosensor applications.[17–21] There has been increased interest recently in the use of lectins in an array format as evidenced by the lectin analysis system of Procognia (Israel) and preliminary work by Angeloni et al. on the deposition of lectins onto a modified dextran surface.[22] For our initial experiments, nine commonly available lectins were used (Table 1). They were arrayed by using a manual arrayer and standard methods on either aldehyde- or epoxide-derivatized glass slides, which yielded spots of ~700 μm in diameter (see Supporting Information).[23,24] A critical modification to the standard methods for protein-array printing was the reduced use of glycerol in the spotting buffers. Glycerol appears to inhibit lectin binding to the glass surface, regardless of the coupling chemistry utilized. In contrast, addition of glycerol to binding buffers after the array had been created had no discernable effect on binding (data not shown). The arrays were interrogated with glycoproteins labeled, by conjugation to lysines, with the fluorescent dye Cy3; thus the protein acts as a labeled handle for the carbohydrate moiety (Figure 1a). Cy3 was chosen due to the wide availability of array scanners with which to detect it. Conjugation of the glycoproteins to Cy3 did not affect lectin binding as observed by dot blot analysis (data not shown).

Table 1. Lectins used in the array and their carbohydrate-binding specificities.

Lectin	Specificity
Canavalia ensiformis (Con A)	branched and terminal mannose, terminal GlcNAc
Galanthus nivalis (GNA)	terminal α-1,3 mannose
Griffonia simplicifolia I (GS-I)	α-galactose
Griffonia simplicifolia II (GS-II)	terminal GlcNAc
Maackia amurensis (MAA)	α-2,3 sialic acid
Glycine Max (SBA)	terminal GalNAc
Sambucus Nigra (SNA)	α-2,6 sialic acid
Ulex europaeus (UEA)	β-fucose
Tritiicum vulgare (WGA)	β-GlcNAc, sialic acid, GalNAc

For our initial analyte we chose ovalbumin, a well-characterized glycoprotein from chicken egg. Ovalbumin has two possible *N*-linked glycosylation sites and is known to have high mannose and/or hybrid *N*-linked glycans.[25] Binding of a 100 μg mL^{-1} solution of Cy3-conjugated ovalbumin to the lectin microarray resulted in a glycopattern characterized by positive signals for WGA, Con A and GS-II (Figure 1). This pattern was observed consistently as demonstrated by two representative arrays. Signals for the positive lectins were well above the background observed with NeutraAvidin (NAv, Fig-

 DOI: 10.1002/cbic.200400403

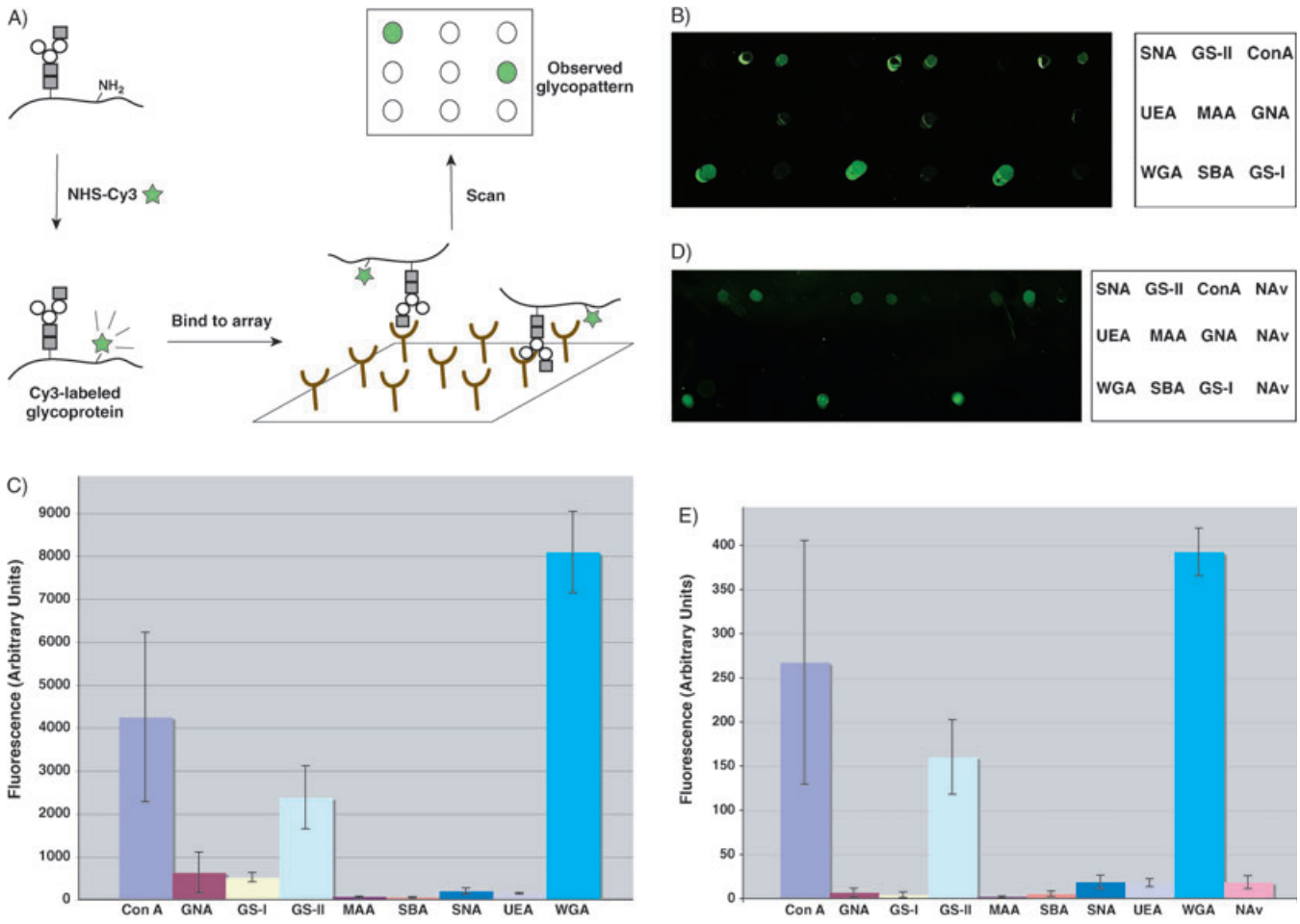

Figure 1. Reproducible glycopatterns are obtained from lectin microarrays on both aldehyde- and epoxide-derivatized slides. A) Schematic representation of the lectin microarray assay. Glycoproteins are labeled with Cy3 by conjugation to lysines by standard methods. The Cy3–glycoproteins are bound to the lectin microarray, and the resulting glycopatterns are analyzed (see Supporting Information for details). B) Cy3-labeled ovalbumin (100 $\mu g\,mL^{-1}$) was used to probe the nine-lectin array on aldehyde-derivatized slides. A portion of the slide with three replicates of the lectin array is shown. C) Graphical representation of the array data from the slide shown in (B). Slides were analyzed by using the Genepix 5.1 software. The average of the median fluorescence signal (arbitrary units) for a single slide is taken. Standard deviations were calculated by using Microsoft Excel. D) Cy3–ovalbumin binding to lectin array on epoxide slides. E) Graphical representation of the array data from the slide shown in (D).

ure 1e), which was used as a negative control. All three lectins bind to GlcNAc, an indicator of hybrid *N*-linked glycans. Con A also binds to mannose; this is consistent with both hybrid and high-mannose-type structures. These data are consistent with those obtained by dot blot analysis of the glycoprotein with biotinylated lectins (Supporting Information). The array signal was found to be linear for all three lectins in the range of 50–300 $\mu g\,mL^{-1}$ of glycoprotein (Supporting Information). Indeed the pattern was still visible at our detection limit of 10 $\mu g\,mL^{-1}$ (data not shown) and was also unchanged when the spot size was decreased to ~80 μm in diameter by using an automated arrayer (Supporting Information). The smaller spot size allows us to potentially fit thousands of spots in a small area, thus enabling us to drastically increase the number of lectins assayed without increasing the amount of analyte. To further probe the specificity of glycan binding to the lectin array, we assayed the ability of monosaccharides to specifically compete with the Cy3–ovalbumin signal (Figure 2). As expected, incubation of the array with the known inhibitory sugar GlcNAc decreased the fluorescence at all three positive lectins to $<10\%$ of the control signal. In contrast, incubation with the noninhibitory sugar galactose had no effect on the fluorescence. This demonstrates the carbohydrate specificity of the binding observed by the lectin microarray.

To characterize the ability of the lectin microarray to yield distinctive glycopatterns, we compared the glycopattern observed with ovalbumin to those of two other Cy3-labeled glycoproteins, bovine submaxillary mucin (BSM) and porcine gastric mucin (PGM). Representative data from slides for each protein, normalized to the highest mean signal on each slide for ease of comparison, are shown in Figure 3. Even with a limited nine-lectin array, unique glycopatterns, reflective of the differences in the glycosylation state of these proteins, were observed. The ovalbumin, which contains only N-linked glycans, as previously described, gave a distinctly different pattern from those of the mucins. Both mucins are known to have fucosylated O-linked glycans, terminal α-GalNAc, GlcNAc, and α-2,6 sialic acid.[26] This correlates well with the positive UEA, GS-I,

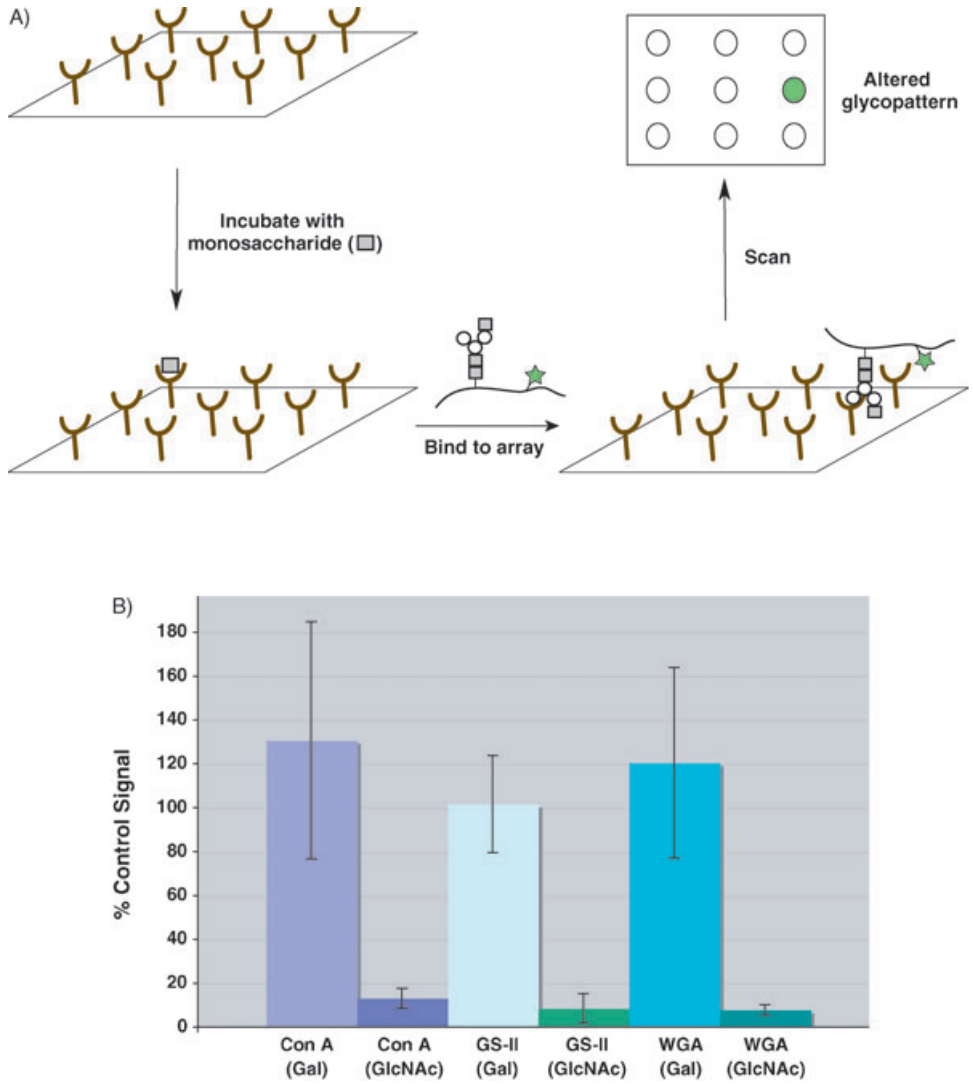

Figure 2. Inhibition of lectin microarray signals with appropriate monosaccharides. A) Slides were incubated with 200 mM of monosaccharide to which the Cy3–glycoprotein sample was then added (final concentration of sugar, 100 mM). After binding, the slides were rinsed and then scanned. B) Effect of the noninhibitory sugar galactose and the inhibitory sugar GlcNAc on lectin signals observed for Cy3–ovalbumin. Signal of lectin as % signal of average median fluorescence for sugar incubated array divided by the average median fluorescence for control lectin array for each positive lectin in the Cy3–ovalbumin array is shown. Error bars were generated by propagation of error. Representative data from a single slide is shown.

SBA, and WGA signals observed (Figure 3). The borderline positive signals for SBA and UEA in the porcine mucin and GS-II in the bovine mucin were confirmed as positive by using monosaccharide inhibition (Supporting Information). Differences between the two mucins can be seen in the GS-II- and WGA-binding levels. The high WGA binding observed is not unexpected as PGM is known to bind to WGA far better than to UEA and provides a second point of difference between the two mucin patterns.[27] The difference in WGA binding between the two mucins most likely reflects the presence of larger, more complex O-linked oligosaccharides on PGM that contain high amounts of galactose and internal GlcNAc compared to the shorter carbohydrate chains observed with BSM.[26] Although both mucins are known to have α-2,6 sialic acids, the lack of observable binding to SNA, which is α-2,6 sialic acid reactive, is not wholly unexpected. Conflicting reports exist about the ability of SNA to bind to α-2,6 sialic acids presented in the context of mucins.[28,29] SNA from E.Y. Laboratories, please explain did not bind significantly to mucins in either our lectin array or in dot blot assays (data not shown). This contrasts with the positive SNA signal for the mucins observed with the DIG kit (Boehringer–Mannheim, data not shown) and suggests that SNA from several sources should be included in future arrays. Hybridization of Cy3-labeled Fetuin, a known SNA-binding protein, to our lectin array gave a strong fluorescence signal for SNA; this indicated that the lectin is still active when bound to the derivatized glass surface (data not shown). The distinct glycopatterns observed even with a minimal nine-lectin array clearly demonstrate the potential of such an array to rapidly pinpoint subtle differences between the glycosylation states of proteins. Indeed, in preliminary experiments, differences between the glycan structures of Cy3- and Cy5-labeled glycoproteins could be determined simultaneously in a mixture (Supporting Information).

Analytical tools for the rapid and systematic analysis of glycosylation are sorely needed to push forward our understanding of the roles of carbohydrates at the biological level. Lectins have been used to gain preliminary information about the glycosylation state of proteins for decades.[18,19] Standard methods for performing lectin-based glycan profiling are useful but constrained by the amounts of sample and analysis time required; this limits the diversity of lectins traditionally used to examine the glycoforms. Although the array demonstrated in this paper consists of only nine lectins, it can easily be extended to include the 62 commercially available lectins, antibodies against carbohydrate epitopes, and lectins from a multitude of sources. Since lectins are the primary means by which cells and bacteria interpret the carbohydrate code, they are capable of binding a diverse array of glycan structures.[30] Thus, an expanded lectin microarray should provide a rapid detailed anal-

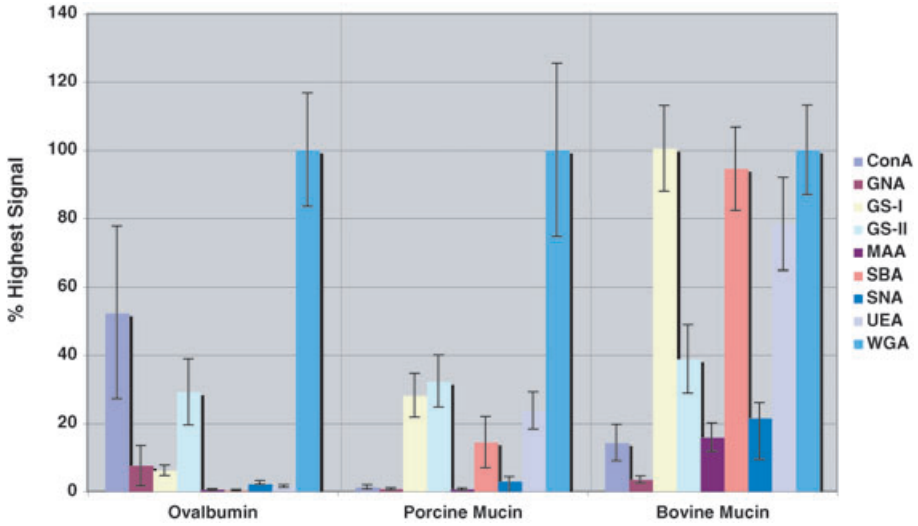

Figure 3. Analysis of multiple glycoproteins reveals distinct glycopatterns. The percentage highest signal was used to allow the patterns to be directly compared. % highest signal = average median spot intensity for a representative slide/average median spot intensity of lectin with the highest signal for that slide (WGA for Porcine Mucin and ovalbumin). The data shown are representative of a minimum of three replicate slides.

ysis of the carbohydrate composition of glycoproteins without utilizing large amounts of sample. In addition, miniaturization of the arrays means that multiple samples could be analyzed simultaneously on a single slide, providing a simple and rapid analytical technique for the profiling of glycosylation states, an essential step in moving forward the field of glycomics. It should be noted that some binding activity may be lost in the coupling of lectins to the array. This is observed in many protein arrays to date and is an ongoing subject of research.[31] The lectin array presented here represents the first step towards the development of microarray methods to rapidly compare cellular glycosylation states, technology critical to the study of carbohydrate function at the systems biology level.

Acknowledgements

We thank Dr. Vishy Ayer, Dr. Andrew D. Ellington, and Jim Collett for their advice and aid, and the University of Texas at Austin and the Arnold and Mabel Beckman Foundation for financial support of this work.

Keywords: analytical methods · carbohydrates · glycomics · microarrays · lectin

[1] A. Varki, *Glycobiology* **1993**, *3*, 97 – 130.
[2] R. Apweiler, H. Hermjakob, N. Sharon, *Biochim. Biophys. Acta* **1999**, *1473*, 4 – 8.
[3] C. R. Bertozzi, L. L. Kiessling, *Science* **2001**, *291*, 2357 – 2364.
[4] J. Hirabayashi, *Trends Biotechnol.* **2003**, *21*, 141 – 143; Discussion p. 143.
[5] S. M. Khersonsky, C. M. Ho, M. A. Garcia, Y. T. Chang, *Curr. Top. Med. Chem.* **2003**, *3*, 617 – 643.
[6] K. R. Love, P. H. Seeberger, *Angew. Chem.* **2002**, *114*, 3733 – 3736; *Angew. Chem. Int. Ed.* **2002**, *41*, 3583 – 3586.
[7] H.-J. Gabius, S. André, H. Kaltner, H. C. Siebert, *Biochim. Biophys. Acta* **2002**, *1572*, 165 – 177.
[8] 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.
[9] M. C. Bryan, F. Fazio, H.-K. Lee, C.-Y. Huang, A. Chang, M. D. Best, D. A. Calarese, O. Blixt, J. C. Paulson, D. Burton, I. A. Wilson, C. H. Wong, *J. Am. Chem. Soc.* **2004**, *126*, 8640 – 8641.
[10] L. Nimrichter, A. Gargir, M. Gortler, R. T. Altstock, A. Shtevi, O. Weisshaus, E. Fire, N. Dotan, R. L. Schnaar, *Glycobiology* **2004**, *14*, 197 – 203.
[11] B. T. Houseman, M. Mrksich, *Chem. Biol.* **2002**, *9*, 443 – 454.
[12] C. Ortiz Mellet, J. M. García Fernández, *ChemBioChem* **2002**, *3*, 819 – 822.
[13] D. Wang, *Proteomics* **2003**, *3*, 2167 – 2175.
[14] E. Pochec, A. Litynska, A. Amoresano, A. Casbarra, *Biochim. Biophys. Acta* **2003**, *1643*, 113 – 123.
[15] G. MacBeath, *Nat. Genet.* **2002**, *32 Suppl*, 526 – 532.
[16] H. Zhu, M. Snyder, *Curr. Opin. Chem. Biol.* **2003**, *7*, 55 – 63.
[17] E. J. M. Van Damme, W. Peumans, A. Pusztai, S. Bardocz, *Handbook of Plant Lectins: Properties and Biomedical Applications*, Wiley, New York, **1998**.
[18] H. Rüdiger, H. J. Gabius, *Glycoconjugate J.* **2001**, *18*, 589 – 613.
[19] I. J. Goldstein, *J. Agric. Food Chem.* **2002**, *50*, 6583 – 6585.
[20] P. Ertl, M. Wagner, E. Corton, S. R. Mikkelsen, *Biosens. Bioelectron.* **2003**, *18*, 907 – 916.
[21] P. Ertl, S. R. Mikkelsen, *Anal. Chem.* **2001**, *73*, 4241 – 4248.
[22] S. Angeloni, J. L. Ridet, N. Kusy, H. Gao, F. Crevoisier, S. Guinchard, S. Kochhar, H. Sigrist, N. Sprenger, *Glycobiology* **2005**, *15*, 31 – 41.
[23] G. MacBeath, S. L. Schreiber, *Science* **2000**, *289*, 1760 – 1763.
[24] B. B. Haab, M. J. Dunham, P. O. Brown, *Adv. Genome Biol.* **2001**, *2*, 1 – 13.
[25] F.-Y. Che, J.-F. Song, X.-X. Shao, K.-Y. Wang, Q.-C. Xia, *J. Chromatogr. A* **1999**, *849*, 599 – 608.
[26] N. G. Karlsson, N. H. Packer, *Anal. Biochem.* **2002**, *305*, 173 – 185.
[27] M. Wirth, K. Gerhardt, C. Wurm, F. Gabor, *J. Controlled Release* **2002**, *79*, 183 – 191.
[28] E. Fischer, R. Brossmer, *Glycoconjugate J.* **1995**, *12*, 707 – 713.
[29] N. Shibuya, I. J. Goldstein, W. F. Broekaert, M. Nsimba-Lubaki, B. Peeters, W. J. Peumans, *J. Biol. Chem.* **1987**, *262*, 1596 – 1601.
[30] H. J. Gabius, H. C. Siebert, S. Andre, J. Jiménez-Barbero, H. Rudiger, *ChemBioChem* **2004**, *5*, 740 – 764.
[31] P. Peluso, D. S. Wilson, D. Do, H. Tran, M. Venkatasubbaiah, D. Quincy, B. Heidecker, K. Poindexter, N. Tolani, M. Phelan, K. Witte, L. S. Jung, P. Wagner, S. Nock, *Anal. Biochem.* **2003**, *312*, 113 – 124.

Received: November 13, 2004
Published online on ■■ ▮▮, 2005

