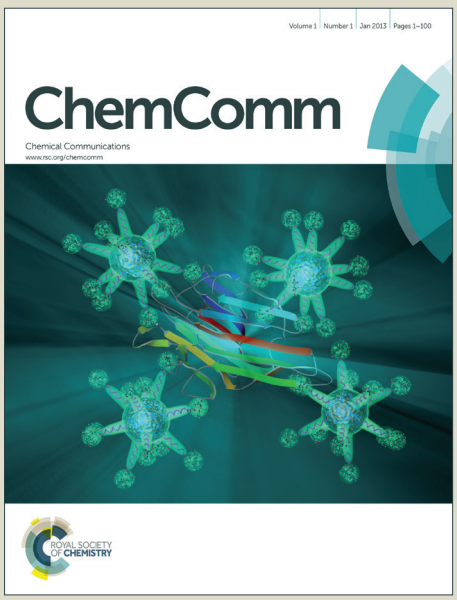
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

ChemComm

RSCPublishing

COMMUNICATION

A multifunctional anomeric linker for the chemoenzymatic synthesis of complex oligosaccharides†

Cite this: DOI: 10.1039/x0xx00000x

Received 00th March 2014,
Accepted 00th Xxxxx 2014

DOI: 10.1039/x0xx00000x

www.rsc.org/chemcomm

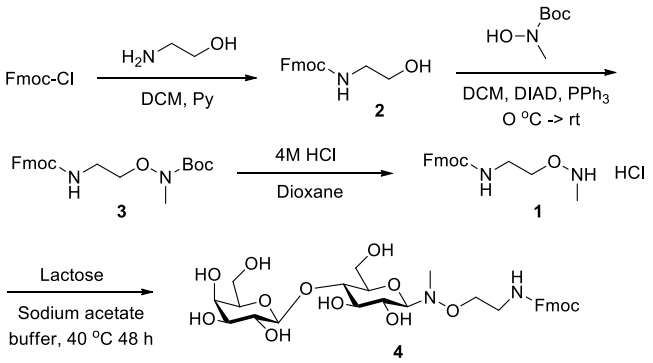
Anthony R. Prudden,[a,b] Zoeisha S. Chinoy,[a,b] Margreet A. Wolfert[a] and Geert-Jan Boons[a,b]*

A new anomeric linker has been developed that facilitates the purification of glycans prepared by chemoenzymatic approaches and can give readily compounds that are appropriately modified for microarray development or glycan derivatives with a free reducing end that are needed as standards for the development of analytical protocols.

Almost all cell surface and secreted proteins are modified by covalently-linked glycans which are essential mediators of biological processes such as protein folding, cell signaling, fertilization, embryogenesis, cell proliferation and tissue morphogenesis.[1] Overwhelming data supports the relevance of glycosylation in pathogen recognition, inflammation, innate immune responses and the development of autoimmune diseases and cancer.[2] The ability of cells to generate information rich glycans has created a new field of research termed "*glycomics*", which seeks to identify and understand the processes involved in the formation of cell type and developmental stage specific oligosaccharide patterns.[3-4] Advances in the robustness and sensitivity of glycan analytics over the past ten years, especially in mass spectrometry-based glycomics,[5] have produced new opportunities for characterizing tissue and cell-specific glycan profiles. Glycan arrays are another key glycomic technology for profiling glycan-protein interactions.[6-9]

Well-defined complex oligosaccharides are critical for the development of contemporary glycomic technologies.[10-12] Such compounds are required for the fabrication of next generation carbohydrate microarrays to examine glycan-protein recognition and to design probes for elucidating pathways of glycoconjugate biosynthesis. Furthermore, well-defined carbohydrates are needed as standards to develop protocols for more accurate identification of glycans in complex mixtures by mass spectrometry, HPLC, capillary electrophoreses or a combination of these methodologies.[13-14] Some glycomic technologies, such as the use of the synthetic glycans as standards for mass spectrometry, require glycans having an unmodified anomeric center. Other uses, such as the development of glycan microarrays, require compounds that have a reactive linker at the anomeric center for immobilization to a surface. A number of attractive approaches have been explored to install an anomeric linker for glycan microarray development using hydrazides,[15-16] hydrazines,[17] hydroxylamines,[18-20] thiosemicarbazides,[21] reductive amination[22] and oxime ligation strategies.[23-24] These approaches take advantage of the free reducing terminus of a glycan and typically these modifications are performed at the end of a synthetic sequence, or to a heterogeneous mixture of glycans extracted from a natural source.

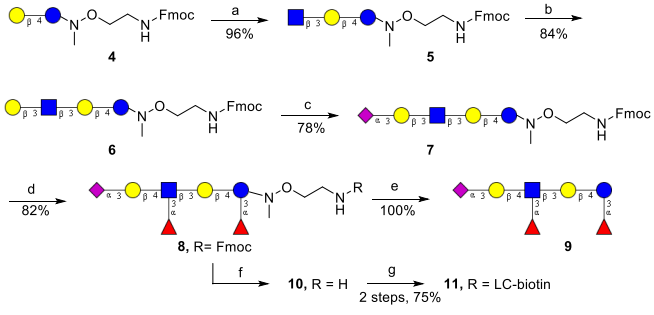
Here, we describe a linker (compound **1**, Scheme 1) that can be installed prior to the enzymatic extensions to facilitate purification of intermediates. Compound **1** contains a *N*-methyl hydroxylamine moiety that is known to react with the anomeric center of reducing sugars to give exclusively β-linked cyclic structures.[18, 20, 25-26] The fluorenylmethyloxycarbonate (Fmoc) group provides a chromophore for UV detection and will endow the compound with sufficient lipophilic character to allow purification by HPLC using a reverse phase column. At the completion of the enzymatic transformations, the Fmoc protecting group can be removed under basic conditions to provide a free amine that can be used for glycan immobilization, aiding in array development. The linker can also be hydrolyzed under mild acidic conditions to give glycans having a free reducing end and such compounds are expected to find use as standards for the development of analytical protocols. Thus, it was anticipated that linker **1** would make it possible to prepare targets for multiple purposes.



Scheme 1 Synthesis of bifunctional linker and lactose conjugation.

Linker **1** was readily prepared by reacting ethanolamine with Fmoc chloride in DCM in the presence of K_2CO_3 followed by treatment with *N*-*t*-butyloxycarbamate-*N*′-methylhydroxylamine under Mitsunobu conditions to provide, after purification by silica gel column chromatography, compound **3**. The butyloxycarbamate (Boc) of **3** was cleaved by treatment with 4 M HCl in dioxane to give **1** in high yield.

To establish the methodology, lactose was modified with the new linker and the resulting compound was converted into derivative **8** (Scheme 2), which contains a sialyl Lewis^x (SLe^x) moiety, a structural motif that is overexpressed by a number of human cancers.[27] Thus, treatment of lactose with **1** (3.5 eq) in a sodium acetate buffer (0.1 M, pH 4.2) at 40 °C for 48 h gave, after purification by preparative reverse phase column chromatography using a mixture of acetonitrile and water as the eluent, compound **4** in a yield of 33%. The equilibrium constants of reactions of *N*-methyl hydroxylamines with carbohydrates are relatively small[26] and therefore these conjugations only proceed to completion when a high concentration of reagents is employed. Due to the relatively low water solubility of **1** such conditions could not be established explaining the modest yield of formation of **4**. In an alternative procedure, we treated lactose with 2-((methylamino)oxy)ethanamine dihydrochloride[19] and the amine of the resulting derivative was protected by Fmoc using a standard procedure to give compound **4** in an overall yield of 87%. In this case, the excellent water solubility of the linker made it possible to use a large excess of the linkering reagent, making it possible to drive the reaction to completion.



Scheme 2 Enzymatic extension of linkered lactose and linker hydrolysis and modification. Reagents and conditions: a) β3GlcNAcT, UDP-GlcNAc; b) GalT-I, UDP-Gal; c) ST3Gal-IV, CMP-Neu5Ac; d) HPα1-3FucT, GDP-Fuc; e) TFA 0.25% *v/v*; f) piperidine 20%/H_2O; g) sulfo-NHS-LC-biotin, PBS buffer, 0.1M, pH 8.0. Glycan symbols: *N*-acetyl-D-glucosamine (GlcNAc, ■); *N*-acetyl neuraminic acid (Neu5Ac, ◆); L-fucose (Fuc, ▲); D-galactose (Gal, ●); and D-glucose (Glu, ●).

Treatment of **4** (Scheme 2) with β1,3-*N*-acetylglucosaminyltransferase (β1,3GlcNAcT),[28] UDP-GlcNAc and Calf Intestine Alkaline Phosphatase (CIAP) resulted in the addition of a β(1,3)-linked GlcNAc moiety to provide trisaccharide **5**. The GlcNAc moiety of **5** was converted into LacNAc by employing β1,4-galactosyltransferase (GalT-I), UDP-Gal and CIAP to give tetrasaccharide **6**. The latter compound was sialylated by ST3Gal-IV, CMP-Neu5Ac and CIAP to give the expected pentasaccharide **7**. Finally, fucosylation of **7** with α1,3-fucosyltransferase (α3FucT)[29] resulted in the modification of the lactose and sialyl-LacNAc moieties to give the target bis-fucosylated derivative **8**. Importantly, after each step, the product could easily be purified by reverse phase column chromatography and characterization of the compounds by NMR and mass spectrometry established structural integrity and homogeneity. The successful preparation of heptasaccharide **8** also demonstrates that the linker is sufficiently stable to undergo multiple enzymatic transformations.

Next, attention was focused on the removal of the linker to give a reducing anomeric center or removal of the Fmoc protecting group to provide a reactive amine, which can be employed for further functionalization. It is well known that anomeric hydroxylamines hydrolyze under acidic conditions,[25-26] however, the challenge was to establish conditions to cleave the hydrophobic linker of the heptasaccharide **8** without affecting the acid labile fucosides and sialoside. Fortunately, treatment of **8** with 0.25% *v/v* TFA in water gave, after an incubation time of 2 h, compound **9** in almost quantitative yield. Alternatively, the Fmoc group of **8** could be removed by treatment with 20% piperidine in water to yield heptasaccharide **10** having an amino-containing anomeric linker. The free amine of **10** was functionalized by reaction with sulfo-NHS-LC-biotin in PBS buffer (0.1 M, pH 8.0) for 3 h to give, after purification by P-2 size exclusion column chromatography, the biotin-modified derivative **11**. Compounds **12-15** (Fig. 1) were prepared in a similar manner and the collection of biotin containing glycans was immobilized on microtiter plates coated with NeutrAvidin. The resulting glycan array was probed with galectin-3, which is a member of an evolutionarily preserved class of protein that has a carbohydrate-recognition domain (CRD) attached to a long *N*-terminal proline- and glycine-rich domain.[30] Galectin-3 has been implicated in a wide range of biological processes including cell adhesion, cell activation and chemo-attraction, cell growth and differentiation and apoptosis. It is also involved in a number of diseases such as cancer, inflammation, fibrosis, heart disease and stroke. Galectin-3 recognizes β-galactosides such as *N*-acetyllactosamine (LacNAc, Galβ1,4GlcNAc), however, it has an extended CRD and can accommodate larger oligosaccharides such as poly-LacNAc and terminal sialyl LacNAc moieties.[31] The interaction of recombinant human galectin-3 with the immobilized glycans was visualized using a rabbit anti-human galectin-3 antibody and secondary detection anti-rabbit IgG antibody labeled with Alexa Fluor 488. Fig. 2 shows significant binding of galectin-3 to

compounds **13** and **14** that have a terminal LacNAc or sialyl-LacNAc moiety, respectively. Compounds **11** and **12**, which contain a terminal Le^x or SLe^x moiety, did not exhibit significant binding, which is in agreement with the ligand requirements of galectin-3.[31-32] These results demonstrate that the novel linkering approach can provide glycans useful for glycomic applications.

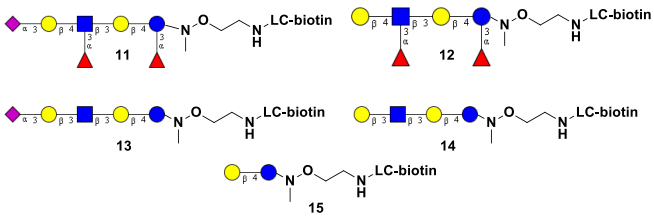


Fig. 1 Library of LC-biotin modified oligosaccharide standards.

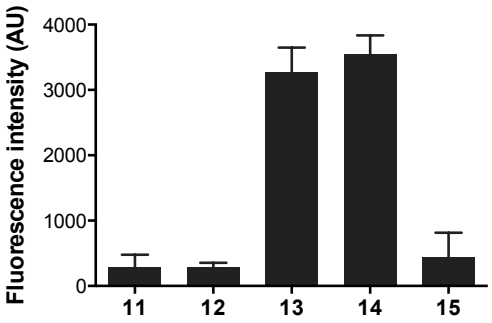
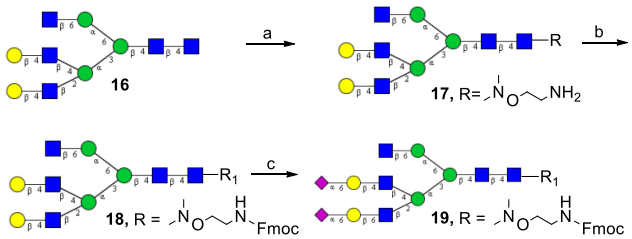


Fig. 2 Galectin-3 binding with target glycans. Biotin-conjugated heparin oligosaccharides **11-15** (10 µM) were allowed to bind for 2 h on NeutrAvidin-coated plates. After washing and incubating with recombinant human Galectin-3 (5 µg/mL; 1 h), rabbit anti-human Galectin-3 antibody (5 µg/mL; 1 h) and Alexa Fluor 488 labeled anti-rabbit IgG (20 µg/mL; 1 h) respectively, fluorescence was read. Data are reported as the means ± SD (n=3).

Recently, we reported a chemoenzymatic strategy for the preparation of libraries of highly complex asymmetrically branched *N*-glycans.[33] Although this methodology can provide the most complex multi-antennary *N*-glycans ever reported, monitoring of the reaction progress and purification of the synthetic intermediates proved to be a stumbling block. To address these difficulties, we examined whether the new linkering methodology could be extended to these types of highly complex glycans. Thus, the previously described decasaccharide **16**[33] was dissolved in a sodium acetate buffer (0.1 M, pH 6.0) and treated with an excess of 2-((methylamino)oxy)ethanamine dihydrochloride salt at 35 °C for 48 h (Scheme 3). The resulting linkered decasaccharide **17** was purified using Sephadex®G-25 superfine size exclusion chromatography. The terminal amine of **17** was protected with an Fmoc functionality using standard procedures, and gratifyingly, the resulting compound **18** could easily be purified by C8 reverse phase column chromatography. The terminal LacNAc moieties of **18** were extended by 2,6-sialosides using ST6Gal-I, CMP-sialic acid and CIAP.[34] It has been reported that ST6Gal-I preferentially modifies the α(1,3)-Man antenna of *N*-linked glycans,[35] and therefore it was not surprising that after a reaction time of 18 h, analysis of the reaction mixture by HPLC using a C8 column, showed a mixture of mono- and bis-sialylated structures (Fig. S2, ESI†). The compound was re-exposed to ST6Gal-I to drive the reaction to completion to give, after purification by C8 reverse phase column chromatography (Fig. S3, ESI†), pure dodecasaccharide **19** in a high yield of 79%. It is important to note that monitoring the progress of enzymatic transformations involving sialic acid by mass spectrometry is challenging due to cleavage of sialic acid residues during the ionization process, and thus, the HPLC-based approach described here provides an attractive alternative.



Scheme 3 Dodecasaccharide linkering and enzymatic extension. Reagents and conditions: a) 2-((methylamino)oxy)ethanamine dihydrochloride, sodium acetate buffer 0.1M, pH 6.0; b) Fmoc-Cl H_2O/ACN; c) ST6Gal-I, CMP-Neu5Ac. Glycan symbols: *N*-acetyl-D-glucosamine (GlcNAc, ■); *N*-acetyl neuraminic acid (Neu5Ac, ◆); D-galactose (Gal, ●); D-glucose (Glu, ●); and D-mannose (Man, ●).

In conclusion, we herein describe a novel multi-functional linker that can readily be attached to the reducing end of unprotected glycans. The lipophilicity of the Fmoc group makes it possible to conveniently purify reaction products by reverse phase column chromatography and monitor reaction progress of enzymatic glycosylations by HPLC. It is to be expected that the use of HILIC column chromatography,[36-37] which exhibits a much higher resolution for glycans than reverse phase based separation approaches, will make it possible to monitor and purify closely related highly complex glycans. The linker of the final products can be cleaved using mild acidic conditions without affecting acid sensitive fucosides or sialic acid residues, to give reducing glycans that can be employed as standards for various analytical purposes. Alternatively, the Fmoc group of the linker can be removed under mild basic conditions to yield a primary amine that can be exploited for glycan array formation. It is the expectation, complex free reducing oligosaccharides isolated from natural sources can also be functionalized by the new linker and therefore it may provide opportunities to fractionate such compounds while preserving the cyclic reducing pyranose structure. *N*-methyl hydroxylamines are stable under glycosidase-catalyzed hydrolysis of glycosides, and therefore it is anticipated that the new linker can also be employed for remodeling of complex carbohydrates using these enzymes.[38] The new linkering approach makes it possible to conveniently prepare complex glycans by chemoenzymatic approaches[39] that can be employed for multiple purposes.

Previously reported linkers do not offer such flexibility for target utilization. In this respect, reductive amination approaches do not allow clean removal of the linker and the resulting linked glycan is relegated to an unnatural open form.[22, 40] Hydrazide-based linkers can form isomeric products,[15-16, 41] which complicate HPLC and

NMR analysis and thus are less attractive for chemo-enzymatic synthesis. Furthermore, such linkers having a pre-installed biotin moiety do not offer the possibility to examine a range of immobilization approaches for microarray development. Previously, we employed 2-((methylamino)oxy)ethanamine to install a linker for glycan microarray development,[33] however the resulting compounds were insufficiently lipophilic and did not contain a chromophore for convenient purification purposes.

This research was supported by the National Institute of General Medical Sciences (R01GM090269 and P01GM107012 to G.-J.B) and the National Cancer Institute (F31CA180478 to A.R.P.) from the US National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. A.R.P. has a Grimes Family Graduate Fellowship in Natural Sciences.

Notes and references

[a] *Complex Carbohydrate Research Center, The University of Georgia, 315 Riverbend Road, Athens, Georgia 30602, USA. E-mail: gjboons@ccrc.uga.edu; Fax: +1 706-542-4412.*
[b] *Department of Chemistry, The University of Georgia, Athens, Georgia, GA 30602, USA.*

† Electronic Supplementary Information (ESI) available: Materials and methods, analytical data and copies of NMR spectra for compounds **1**-**15**, **18** and **6S**. See DOI: 10.1039/c000000x/

1 G. W. Hart and R. J. Copeland, *Cell,* 2010, **143**, 672-676.
2 K. Ohtsubo and J. D. Marth, *Cell,* 2006, **126**, 855-867.
3 J. C. Paulson, O. Blixt and B. E. Collins, *Nat. Chem. Biol.,* 2006, **2**, 238-248.
4 R. D. Cummings and J. M. Pierce, *Chem. Biol.,* 2014, **21**, 1-15.
5 S. J. North, P. G. Hitchen, S. M. Haslam and A. Dell, *Curr. Opin. Struct. Biol.,* 2009, **19**, 498-506.
6 N. Laurent, J. Voglmeir and S. L. Flitsch, *Chem. Commun.,* 2008, 4400-4412.
7 O. Oyelaran and J. C. Gildersleeve, *Curr. Opin. Chem. Biol.,* 2009, **13**, 406-413.
8 D. F. Smith, X. Song and R. D. Cummings, *Methods Enzymol.,* 2010, **480**, 417-444.
9 C. D. Rillahan and J. C. Paulson, *Annu. Rev. Biochem.,* 2011, **80**, 797-823.
10 T. J. Boltje, T. Buskas and G. J. Boons, *Nat. Chem.,* 2009, **1**, 611-622.
11 L. L. Kiessling and R. A. Splain, *Annu. Rev. Biochem.,* 2010, **79**, 619-653.
12 B. Lepenies, J. Yin and P. H. Seeberger, *Curr. Opin. Chem. Biol.,* 2010, **14**, 404-411.
13 D. J. Harvey, *J. Chromatogr. B,* 2011, **879**, 1196-1225.
14 W. R. Alley, Jr., B. F. Mann and M. V. Novotny, *Chem. Rev.,* 2013, **113**, 2668-2732.
15 E. A. Bayer, H. Benhur and M. Wilchek, *Anal. Biochem.,* 1988, **170**, 271-281.
16 C. Leteux, R. A. Childs, W. G. Chai, M. S. Stoll, H. Kogelberg and T. Feizi, *Glycobiology,* 1998, **8**, 227-236.
17 J. Auge and N. Lubin-Germain, *J. Carbohydr. Chem.,* 2000, **19**, 379-392.
18 F. Peri, P. Dumy and M. Mutter, *Tetrahedron,* 1998, **54**, 12269-12278.
19 O. Bohorov, H. Andersson-Sand, J. Hoffmann and O. Blixt, *Glycobiology,* 2006, **16**, 21C-27C.
20 E. Clo, O. Blixt and K. J. Jensen, *Eur. J. Org. Chem.,* 2010, 540-554.
21 E. C. Rodriguez, L. A. Marcaurelle and C. R. Bertozzi, *J. Org. Chem.,* 1998, **63**, 7134-7135.
22 X. Z. Song, B. Y. Xia, S. R. Stowell, Y. Lasanajak, D. F. Smith and R. D. Cummings, *Chem. Biol.,* 2009, **16**, 36-47.
23 Y. Liu, T. Feizi, M. A. Campanero-Rhodes, R. A. Childs, Y. Zhang, B. Mulloy, P. G. Evans, H. M. Osborn, D. Otto, P. R. Crocker and W. Chai, *Chem. Biol.,* 2007, **14**, 847-859.
24 H. Shimaoka, H. Kuramoto, J. Furukawa, Y. Miura, M. Kurogochi, Y. Kita, H. Hinou, Y. Shinohara and S. I. Nishimura, *Chem.-Eur. J.,* 2007, **13**, 1664-1673.
25 J. Kalia and R. T. Raines, *Angew. Chem. Int. Ed.,* 2008, **47**, 7523-7526.
26 A. V. Gudmundsdottir, C. E. Paul and M. Nitz, *Carbohydr. Res.,* 2009, **344**, 278-284.
27 J. L. Magnani, *Glycobiology,* 1991, **1**, 318-320.
28 B. Sauerzapfe, K. Krenek, J. Schmiedel, W. W. Wakarchuk, H. Pelantova, V. Kren and L. Elling, *Glycoconjugate J.,* 2009, **26**, 141-159.
29 W. Wang, T. Hu, P. A. Frantom, T. Zheng, B. Gerwe, D. S. Del Amo, S. Garret, R. D. Seidel, III and P. Wu, *Proc. Natl. Acad. Sci. U. S. A.,* 2009, **106**, 16096-16101.
30 J. Dumic, S. Dabelic and M. Flogel, *Biochim. Biophys. Acta, Gen. Subjects,* 2006, **1760**, 616-635.
31 S. R. Stowell, C. M. Arthur, P. Mehta, K. A. Slanina, O. Blixt, H. Leffler, D. F. Smith and R. D. Cummings, *J. Biol. Chem.,* 2008, **283**, 10109-10123.
32 J. Hirabayashi, T. Hashidate, Y. Arata, N. Nishi, T. Nakamura, M. Hirashima, T. Urashima, T. Oka, M. Futai, W. E. G. Muller, F. Yagi and K. Kasai, *Biochim. Biophys. Acta, Gen. Subjects,* 2002, **1572**, 232-254.
33 Z. Wang, Z. S. Chinoy, S. G. Ambre, W. Peng, R. McBride, R. P. de Vries, J. Glushka, J. C. Paulson and G. J. Boons, *Science,* 2013, **341**, 379-383.
34 Y. Ichikawa, J. L. C. Liu, G. J. Shen and C. H. Wong, *J. Am. Chem. Soc.,* 1991, **113**, 6300-6302.
35 A. W. Barb, E. K. Brady and J. H. Prestegard, *Biochemistry,* 2009, **48**, 9705-9707.
36 M. Wuhrer, A. R. de Boer and A. M. Deelder, *Mass Spectrom. Rev.,* 2009, **28**, 192-206.
37 G. Zauner, A. M. Deelder and M. Wuhrer, *Electrophoresis,* 2011, **32**, 3456-3466.
38 A. Iqbal, H. Chibli and C. J. Hamilton, *Carbohydr. Res.,* 2013, **377**, 1-3.
39 R. M. Schmaltz, S. R. Hanson and C. H. Wong, *Chem. Rev.,* 2011, **111**, 4259-4307.
40 X. Song, B. A. Johns, H. Ju, Y. Lasanajak, C. Zhao, D. F. Smith and R. D. Cummings, *ACS Chem. Biol.,* 2013, **8**, 2478-2483.
41 B. Bendiak, *Carbohydr. Res.,* 1997, **304**, 85-90.

