JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Chemoenzymatic Synthesis of *O*-Mannosylpeptides in Solution and on Solid Phase

Robert Šardzík,[†,§] Anthony P. Green,[†,§] Nicolas Laurent,[†] Peter Both,[†] Carolina Fontana,[‡] Josef Voglmeir,[†] Martin J. Weissenborn,[†] Rose Haddoub,[†] Paola Grassi,[#] Stuart M. Haslam,[#] Göran Widmalm,[‡] and Sabine L. Flitsch*[,†]

[†]School of Chemistry, Manchester Interdisciplinary Biocentre, The University of Manchester, Manchester M1 7DN, U.K.
[‡]Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden
[#]Division of Molecular Biosciences, Faculty of Natural Sciences, Imperial College London, London, SW7 2AZ, U.K.

Supporting Information

ABSTRACT: *O*-Mannosyl glycans are known to play an important role in regulating the function of α-dystroglycan (α-DG), as defective glycosylation is associated with various phenotypes of congenital muscular dystrophy. Despite the well-established biological significance of these glycans, questions regarding their precise molecular function remain unanswered. Further biological investigation will require synthetic methods for the generation of pure samples of homogeneous glycopeptides with diverse sequences. Here we describe the first total syntheses of glycopeptides containing the tetrasaccharide NeuNAcα2-3Galβ1-4GlcNAcβ1-2Manα, which is reported to be the most abundant *O*-mannosyl glycan on α-DG. Our approach is based on biomimetic stepwise assembly from the reducing end and also gives access to the naturally occurring mono-, di-, and trisaccharide substructures. In addition to the total synthesis, we have developed a "one-pot" enzymatic cascade leading to the rapid synthesis of the target tetrasaccharide. Finally, solid-phase synthesis of the desired glycopeptides directly on a gold microarray platform is described.

It is predicted that over 50% of proteins in the human body are glycosylated.[1] These complex glycans are known to play a critical role in the regulation of a diverse range of biological processes.[2] However, despite recent advances in the field of glycomics, the precise function of these carbohydrates is poorly understood, largely due to difficulties in obtaining pure samples of homogeneous glycopeptides. Isolation of significant quantities of these structures from natural sources is extremely challenging. As a result, the development of new methodologies for the efficient synthesis of well-defined glycopeptides is of great interest. Several elegant strategies for the synthesis of peptides containing both *N*- and *O*-linked glycans have been reported.[3] *O*-Mannosyl peptides are an important class of structures distinct from the usual *O*-glycans, but have thus far only been identified on one human protein, α-dystroglycan (α-DG), a heavily glycosylated protein found in muscle and brain tissue. Here we report the first total synthesis of glycopeptides containing an *O*-mannnosyl tetrasaccharide, which represents the textbook example of this class of biomolecule.[4]

α-DG is a cell surface glycoprotein which acts as a receptor for both extracellular matrix proteins[5,6] and a number of arenaviruses, including Lymphocytic choriomeningitis virus (LCMV) and Lassa fever virus (LFV).[7] In order to conserve its biological functions, α-DG requires extensive post-translational glycosylation. In addition to typical *N*-glycosylation and mucin-type *O*-GalNAc glycans, a number of *O*-mannosyl glycans have been detected and characterized.[8−10] This mode of post-translational modification has rarely been described in mammals, but has an important role in the function of α-DG, since defects in *O*-mannosyl glycosylation lead to loss of extracellular ligand binding activity, resulting in various phenotypes of congenital muscular dystrophy (CMD).[11] To date, mutations in six genes which encode for actual or putative glycosyltransferases have been identified in patients with various forms of CMD.[12−17] The molecular functions of two of these putative glycosyltransferases, fukutin and fukutin-related protein (FKRP), remain unclear. It has recently been reported that like-acetylglucosaminyltransferase (LARGE) is responsible for a post-phosphoryl modification of a phosphorylated *O*-mannosyl glycan.[18] An impressive total synthesis of this structure has recently been achieved using a traditional chemical approach.[19] Protein *O*-mannosyl transferase 1 (POMT1), POMT2, and protein *O*-mannose β-1,2-*N*-acetylglucosaminyltransferase 1 (POMGnT1) are known to catalyze the first two steps in the biosynthesis of the *O*-linked tetrasaccharide NeuNAcα2-3Galβ1-4GlcNAcβ1-2Manα (Figure 1).[20,14] This tetrasaccharide has been reported to be the most abundant *O*-mannosyl glycan on α-DG from a range of different species and tissues, suggesting its relevance to the basic functional role of this glycoprotein.[9,10,21]

The binding of glycopeptides to their epitopes is often highly dependent upon peptide sequence as well as glycan structure. As a result, biological studies to illucidate the role of the *O*-mannosyl glycans found on α-DG will require the synthesis of well-defined glycopeptides with diverse sequences. Although two syntheses of the fully assembled tetrasaccharide linked to a serine/threonine residue have been described previously,[22,23] the synthesis of the biologically relevant glycopeptides has not been reported. We envisioned an approach to the synthesis of

Received: December 20, 2011
Published: February 28, 2012

dx.doi.org/10.1021/ja211861m | *J. Am. Chem. Soc.* 2012, 134, 4521−4524

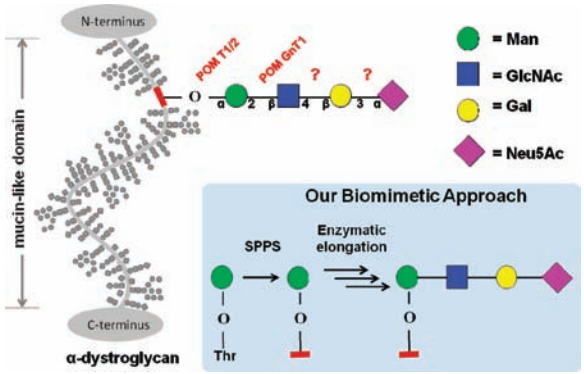
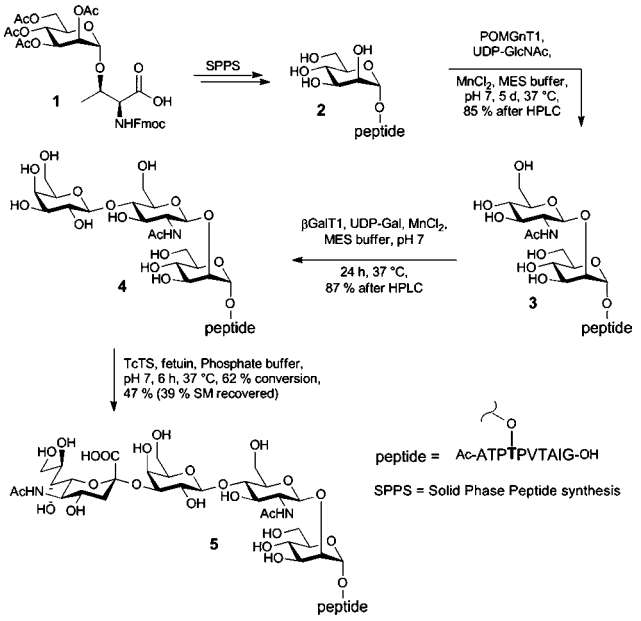


Figure 1. α-DG has a mucin-like domain containing the tetrasaccharide NeuNAcα2-3Galβ1-4GlcNAcβ1-2Manα. The first two glycosidic linkages are formed by a POMT1/POMT2 complex and POMGnT1, respectively. The enzymes responsible for the final two steps have yet to be defined. CFG nomenclature is used to represent the glycan structures.

these structures which closely mimics their biosynthesis. Our approach commences with the chemical synthesis of mannosyl peptides (Figure 1). The installation of the α-linked mannose prior to peptide synthesis is synthetically straightforward and avoids the need for POMT1/POMT2 glycosyltransferase catalysts and the expensive sugar donor mannosylphosphoryldolichol. Subsequent use of three consecutive, enzymatic glycosylation reactions to synthesize the tetrasaccharide would lead to a very short, flexible, and efficient synthetic route mimicking the putative biosynthetic pathway.

The synthesis commenced with the known manno-threonine building block **1** (Scheme 1).[24] This building block was incorporated into standard Fmoc-based SPPS (Supporting Information (SI)). After cleavage from the resin, deprotection of the per-acetylated mannose moiety was achieved using NaOMe/MeOH (pH 10) to yield manno-peptide **2**, containing a natural peptide sequence of α-DG (amino acid residues 317–326). The GlcNAc moiety was then introduced with the desired β1,2 linkage using human POMGnT1, the enzyme naturally responsible for the attachment of a GlcNAc residue onto *O*-mannosylated protein in the α-DG biosynthesis. POMGnT1 was heterologously expressed in *Pichia pastoris* using a protocol recently developed in our laboratory (SI). The reaction was regularly monitored by HPLC and was shown to be complete in 5 days on a 50 mg scale, providing the disaccharide **3** in 85% yield after HPLC purification. It is important to note that the rate of this reaction can be greatly enhanced on a smaller scale by increasing the concentration of enzyme. The manno-threonine fragment **1** can be readily incorporated into SPPS to produce a range of natural and non-natural manno-peptide sequences. Interestingly, we have recently observed that the activity of POMGnT1 toward a particular substrate is highly dependent upon peptide sequence.[25] Further studies may reveal the minimum sequence requirements to maintain enzyme activity.

Scheme 1. Total Synthesis of Glycopeptide 5, Containing *O*-Mannosyl Glycan NeuNAcα2-3Galβ1-4GlcNAcβ1-2Manα



The enzymes responsible for catalyzing the final two steps of the biosynthesis of the tetrasaccharide have yet to be defined. Instead, enzymes were recruited from alternative biosynthetic pathways to complete the synthesis. Bovine β1,4-galactosyltransferase (β1,4-GalT, EC 2.4.1.38) is well characterized and has been extensively used for chemoenzymatic synthesis of glycoconjugates owing to its broad substrate specificity.[26] Using this enzyme, introduction of the galactose with the required β1,4 linkage was achieved after 24 h of incubation at 37 °C, yielding trisaccharide **4** in 87% yield. Finally, a *trans*-sialidase from *Trypanozoma cruzi* (TcTS) was used for attachment of the terminal sialic acid with the required α2,3 configuration.[27] Contrary to sialyltransferases, the TcTS does not require expensive CMP-NeuNAc donor and is able to catalyze *trans*-glycosylation from the sialoprotein fetuin. However, careful monitoring of the reaction is required, as hydrolysis of the newly formed sialosidic linkage can occur over prolonged reaction times. As a result, these reactions are difficult to drive to completion, and a mixture of starting material and product is generally produced. Nevertheless, after 6 h incubation at 37 °C, the target glycopeptide **5** was successfully formed in 47% yield, along with 39% recovered trisaccharide.

The use of three sequential enzymatic elongation steps allows rapid assembly of the tetrasaccharide unit in a manner not possible using traditional chemical glycan synthesis, which is plagued with protecting group manipulations and selectivity issues. This synthesis has enabled production of the final tetrasaccharide in milligram quantities, allowing us to unambiguously characterize the position and stereochemistry of the glycan linkages, as well as the peptide sequence, using detailed NMR studies. The sequential addition of sugar residues is evident from the anomeric region of the ^{13}C NMR spectra (Figure 2). The complete ^{1}H and ^{13}C NMR assignments of **3**–**5** (SI) were facilitated by first predicting the chemical shifts of the oligosaccharide-threonine structure by the computer program CASPER[28] and subsequently analyzing the 2D NMR spectra. For sequential information on the peptide structure, a BS-CT-HMBC experiment[29] proved highly informative as residues could be linked together via heteronuclear alternating intra-residue two-bond and inter-residue three-bond correlations. Further evidence of the tetrasaccharide structure was provided by extensive mass spectrometric analysis (SI). An additional advantage of our synthetic strategy is that it provides simple access to the corresponding mono-, di-, and trisaccharide-containing intermediates, which are frequently detected on α-DG due to

dx.doi.org/10.1021/ja211861m | J. Am. Chem. Soc. 2012, 134, 4521–4524

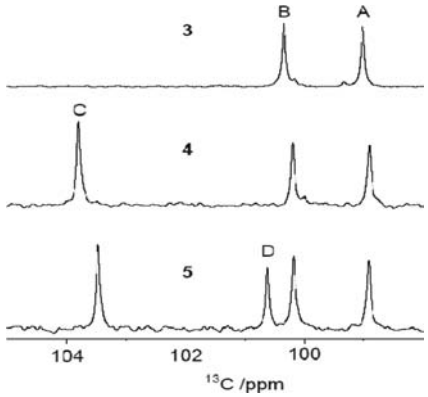


Figure 2. Anomeric region of ^{13}C NMR spectra of disaccharide-peptide **3** (top), trisaccharide-peptide **4** (middle), and tetrasaccharide-peptide **5** (bottom). Resonances are denoted by capital letters: (A) αMan, (B) βGlcNAc, (C) βGal, and (D) NeuNAc.

heterogeneous glycosylation. Our synthesis has provided samples of the final tetrasaccharide **5** as well as intermediates **2**, **3**, and **4** as standards for NMR and glycomic databases. This may simplify the identification of *O*-mannosyl glycans on mammalian proteins other than α-DG, for example, in the gastrointestinal tract of mice.[30]

The use of "one-pot", multiple glycosylations represents an attractive method for the synthesis of complex glycans.[31] Taking advantage of the highly selective nature of the enzymes involved in our synthesis, we have developed a "one-pot" approach to the assembly of glycopeptide **5**. Manno-peptide **2** (1 mg) was added to a premixed solution containing the three enzymes and the three sugar donors (UDP-GlcNAc, UDP-Gal, and fetuin), and the reaction was monitored by HPLC (Figure 3). Following incubation at 37 °C for 24 h, peaks

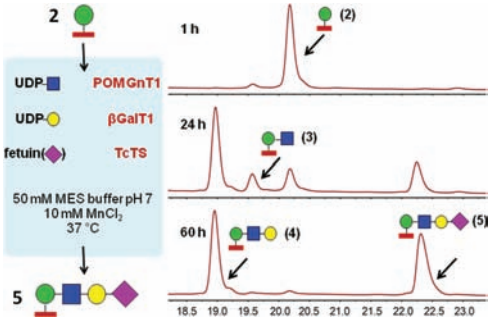


Figure 3. HPLC profiles monitoring the "one-pot" conversion of manno-peptide **2** to tetrasaccharide **5**.

corresponding to glycopeptides **2**–**5** were observed. After 60 h, only traces of the mono- and disaccharide were detected, and the anticipated tri- and tetrasaccharides were formed with good conversion as a 50:50 mixture.[32] This "one-pot" enzymatic cascade significantly reduces the time required to produce the desired tetrasaccharide and alleviates the need for the intermediate purification steps. This methodology should now enable the rapid assembly of this tetrasaccharide on a range of natural peptide sequences of α-DG.

Having developed an efficient solution synthesis, we turned our attention to preparation of these *O*-mannosyl glycopeptides on a solid platform. In recent years, carbohydrate microarrays have emerged as powerful tools to study glycoenzyme specificity and investigate the interactions of carbohydrates with a wide range of binding partners, including proteins, viruses, and whole cells, leading to significant advances in the field of glycomics.[33] Despite the numerous advantages of these glycan microarrays, their production is severely hampered by difficulties in obtaining pure samples of well-defined carbohydrates through either synthesis or isolation. Additionally, isolated carbohydrates must be further modified with a linker to allow attachment to the array surface. One approach to overcome this problem involves solid-phase synthesis of the required glycan directly on the microarray surface.[34] We now report solid-phase synthesis of the *O*-glycan NeuNAcα2-3Galβ1-4GlcNAcβ1-2Manα, attached to a natural peptide sequence of α-DG (amino acid residues 373–384), immobilized on a gold platform. The first step was the formation of an *N*-hydroxysuccinimide (NHS)-functionalized self-assembled monolayer (SAM) (Figure 4). SAMs of alkanethiols on gold

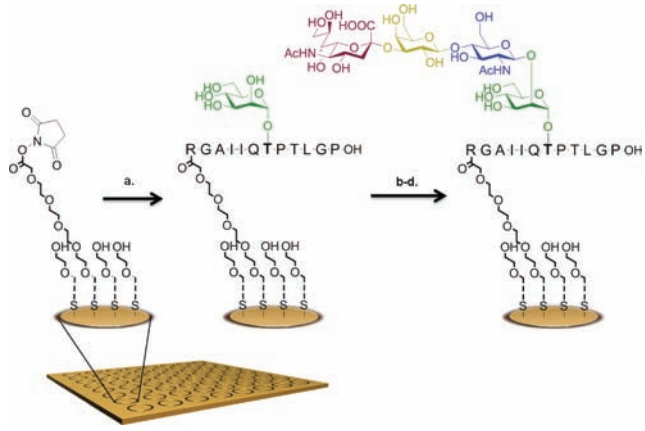


Figure 4. Solid-phase synthesis of target glycopeptide on a gold platform. Conditions: (a) RGAIIQT(Man)PTLGPOH, RT; (b) POMGnT1, UDP-GlcNAc, $MnCl_2$, MES buffer, pH 7, 37 °C; (c) βGalT1, UDP-Gal, $MnCl_2$, MES buffer, pH 7, 37 °C; (d) TcTs, fetuin, phosphate buffer, pH 7, 37 °C.

surfaces provide well-established platforms for carbohydrate microarrays.[35] Attachment of the chemically synthesized mannopeptide was achieved through the amine at the N-terminus of the peptide by formation of an amide bond. The three enzymatic elongation steps described previously (see Scheme 1) were then successfully carried out on the immobilized glycopeptide to produce the desired tetrasaccharide attached to the gold surface. Each glycosylation step, as well as the attachment of the mannopeptide to the surface, was conveniently monitored using MALDI-TOF MS.[36] The first two enzymatic steps proceeded to completion, providing access to uniform monolayers of the di- and trisaccharide structures. Incomplete conversion in the final step led to a mixture of glycoforms on the gold surface, mimicking the biological heterogeneity encountered on the surface of cells. The use of POMGnT1 directly on a microarray platform is of particular interest since this technology will allow the substrate specificity of this important human enzyme to be probed. This solid-phase synthesis has a number of advantages over solution-phase synthesis: Only minute quantities of material are required, minimizing the use of valuable enzymes and mannopeptides. HPLC purification steps associated with solution-phase synthesis are avoided since non-covalently bound reagents and enzymes can be simply washed from the surface. This technology, coupled with our recent studies describing "spot

dx.doi.org/10.1021/ja211861m | *J. Am. Chem. Soc.* 2012, 134, 4521–4524

synthesis" of peptides directly on the array,[37] has the potential to provide rapid access to the tetrasaccharide and its truncated intermediates attached to a diverse range of α-DG peptide sequences. These structures can be produced in parallel on the array in a high-throughput manner to facilitate biological studies.

In summary, we have developed the first syntheses of glycopeptides containing the *O*-mannosyl glycan NeuNAcα2-3Galβ1-4GlcNAcβ1-2Manα. The synthesis of the tetrasaccharide fragment was achieved in a highly efficient manner using three consecutive enzymatic glycosylations. We have demonstrated these reactions both in a "one-pot" fashion in solution and on solid phase, providing rapid access to the desired structures and their intermediates. This technology is currently being used to produce a library of glycopeptides to investigate the role of this unusual glycan in the binding of α-DG to its various receptors and to provide standards for NMR and glycomic databases.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Experimental procedures, characterization of new compounds, and complete refs 12−15. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

sabine.flitsch@manchester.ac.uk

Author Contributions

[§]These authors contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by grants from the EPSRC, the BBSRC, the Royal Society (Wolfson Award to S.L.F.), the Knut and Alice Wallenberg foundation, the Swedish Research Council, and the European commission.

■ REFERENCES

(1) Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochim. Biophys. Acta, Gen. Subj.* **1999**, *1473*, 4.

(2) (a) Spiro, R. G. *Glycobiology* **2002**, *12*, 43R. (b) Paulson, J. C. *Trends Biochem. Sci.* **1989**, *14*, 272. (c) Hart, G. W.; Copeland, R. J. *Cell* **2010**, *143*, 672.

(3) (a) Brocke, C.; Kunz, H. *Bioorg. Med. Chem.* **2002**, *10*, 3085. (b) Pratt, M. R.; Bertozzi, C. R. *Chem. Soc. Rev.* **2005**, *34*, 58. (c) Davis, B. G. *Chem. Rev.* **2002**, *102*, 579. (d) Arsequell, G.; Valencia, G. *Tetrahedron: Asymmetry* **1999**, *10*, 3045. (e) George, S. K.; Schwientek, T.; Holm, B. R.; Reis, C. A.; Clausen, H.; Kihlberg, J. *J. Am. Chem. Soc.* **2001**, *123*, 11117. (f) Mezzato, S.; Unverzagt, C. *Carbohydr. Res.* **2010**, *345*, 1306.

(4) Varki, A.; Sharon, N. In *Essentials of Glycobiolgy*, 2nd ed.; Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2008; p 14.

(5) Ibraghimov-Beskrovnaya, O.; Ervasti, J. M.; Leveille, C. J.; Slaughter, C. A.; Sernett, S. W.; Campbell, K. P. *Nature* **1992**, *355*, 696.

(6) Gee, S. H.; Montanaro, F.; Lindenbaum, M. H.; Carbonetto, S. *Cell* **1994**, *77*, 675.

(7) Cao, W.; Henry, M. D.; Borrow, P.; Yamada, H.; Elder, J. H.; Ravkov, E. V.; Nichol, S. T.; Compans, R. W.; Campbell, K. P.; Oldstone, M. B. A. *Science* **1998**, *282*, 2079.

(8) Chiba, A.; Matsumura, K.; Yamada, H.; Inazu, T.; Shimizu, T.; Kusunoki, S.; Kanazawa, I.; Kobata, A.; Endo, T. *J. Biol. Chem.* **1997**, *272*, 2156.

(9) Nilsson, J.; Nilsson, J.; Larson, G. R.; Grahn, A. *Glycobiology* **2010**, *20*, 1160.

(10) Stalnaker, S. H.; Hashmi, S.; Lim, J.-M.; Aoki, K.; Porterfield, M.; Gutierrez-Sanchez, G.; Wheeler, J.; Ervasti, J. M.; Bergmann, C.; Tiemeyer, M.; Wells, L. *J. Biol. Chem.* **2010**, *285*, 24882.

(11) Michele, D. E.; K. P. Campbell, K. P. *J. Biol. Chem.* **2003**, *278*, 15457.

(12) Beltrán-Valero de Bernabé, D.; et al. *Am. J. Hum. Genet.* **2002**, *71*, 1033.

(13) van Reeuwijk, J.; et al. *J. Med. Genet.* **2005**, *42*, 907.

(14) Yoshida, A.; et al. *Dev. Cell* **2001**, *1*, 717.

(15) Kobayashi, K.; et al. *Nature* **1998**, *394*, 388.

(16) Brockington, M.; Blake, D. J.; Prandini, P.; Brown, S. C.; Torelli, S.; Benson, M. A.; Ponting, C. P.; Estournet, B.; Romero, N. B.; Mercuri, E.; Voit, T.; Sewry, C. A.; Guicheney, P.; Muntoni, F. *Am. J. Hum. Genet.* **2001**, *69*, 1198.

(17) Longman, C.; Brockington, M.; Torelli, S.; Jimenez-Mallebrera, C.; Kennedy, C.; Khalil, N.; Feng, L.; Saran, R. K.; Voit, T.; Merlini, L.; Sewry, C. A.; Brown, S. C.; Muntoni, F. *Hum. Mol. Genet.* **2003**, *12*, 2853.

(18) Yoshida-Moriguchi, T.; Yu, L.; Stalnaker, S. H.; Davis, S.; Kunz, S.; Madson, M.; Oldstone, M. B. A.; Schachter, H.; Wells, L.; Campbell, K. P. *Science* **2010**, *327*, 88.

(19) Mo, K.-F.; Fang, T.; Stalnaker, S. H.; Kirby, P. S.; Liu, M.; Wells, L.; Pierce, M.; Live, D. H.; Boons, G.-J. *J. Am. Chem. Soc.* **2011**, *133*, 14418.

(20) Manya, H.; Chiba, A.; Yoshida, A.; Wang, X.; Chiba, Y.; Jigami, Y.; Margolis, R. U.; Endo, T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 500.

(21) Sasaki, T.; Yamada, H.; Matsumura, K.; Shimizu, T.; Kobata, A.; Endo, T. *Biochim. Biophys. Acta, Gen. Subj.* **1998**, *1425*, 599.

(22) Matsuo, I.; Isomura, M.; Ajisaka, K. *Tetrahedron Lett.* **1999**, *40*, 5047.

(23) (a) Seifert, J.; Ogawa, T.; Ito, Y. *Tetrahedron Lett.* **1999**, *40*, 6803. (b) Seifert, J.; Ogawa, T.; Kurono, S.; Ito, Y. *Glycoconjugate J.* **2000**, *17*, 407.

(24) Varon, D.; Lioy, E.; Patarroyo, M. E.; Schratt, X.; Unverzagt, C. *Aust. J. Chem.* **2002**, *55*, 161.

(25) Voglmeir, J.; Kaloo, S.; Laurent, N.; Meloni, M. M.; Bohlmann, L.; Wilson, I. B. H.; Flitsch, S. L. *Biochem. J.* **2011**, *436*, 447.

(26) Palcic, M. M. *Curr. Opin. Biotechnol.* **1999**, *10*, 616.

(27) Ferrero-García, M. A.; Trombetta, S. E.; Sánchez, D. O.; Reglero, A.; Frasch, A. C. C.; Parodi, A. J. *Eur. J. Biochem.* **1993**, *213*, 765.

(28) Lundborg, M.; Widmalm, G. *Anal. Chem.* **2011**, *83*, 1514.

(29) Claridge, T. D. W.; Pérez-Victoria, I. *Org. Biomol. Chem.* **2003**, *1*, 3632.

(30) Ismail, M. N.; Stone, E. L.; Panico, M.; Lee, S. H.; Luu, Y.; Ramirez, K.; Ho, S. B.; Fukuda, M.; Marth, J. D.; Haslam, S. M.; Dell, A. *Glycobiology* **2011**, *21*, 82.

(31) (a) Bézay, N.; Dudziak, G.; Liese, A.; Kunz, H. *Angew. Chem., Int. Ed.* **2001**, *40*, 2292. (b) Dudziak, G.; Bézay, N.; Schwientek, T.; Clausen, H.; Kunz, H.; Liese, A. *Tetrahedron* **2000**, *56*, 5865.

(32) After 48 h the tri- and tetrasaccharide were formed as a 65:35 mixture. At this stage additional fetuin was added to increase the yield of the desired tetrasaccharide (SI).

(33) (a) Horlacher, T.; Seeberger, P. H. *Chem. Soc. Rev.* **2008**, *37*, 1414. (b) Feizi, T.; Fazio, F.; Chai, W.; Wong, C.-H. *Curr. Opin. Struct. Biol.* **2003**, *13*, 637.

(34) (a) Laurent, N.; Haddoub, R.; Flitsch, S. L. *Trends Biotechnol.* **2008**, *26*, 328. (b) Fazio, F.; Bryan, M. C.; Blixt, O.; Paulson, J. C.; Wong, C.-H. *J. Am. Chem. Soc.* **2002**, *124*, 14397.

(35) Houseman, B. T.; Mrksich, M. *Chem. Biol.* **2002**, *9*, 443.

(36) Su, J.; Mrksich, M. *Angew. Chem., Int. Ed.* **2002**, *41*, 4715.

(37) Laurent, N.; Haddoub, R.; Voglmeir, J.; Wong, S. C. C.; Gaskell, S. J.; Flitsch, S. L. *ChemBioChem* **2008**, *9*, 2592.