An Efficient Synthesis of a Dimer of the Tetrasaccharide Present in Motif B of the *Mycobacterium tuberculosis* Cell Wall

Xiangdong Mei,[a] Jun Ning*[a,b]

[a] Research Center for Eco-Environmental Sciences, Academia Sinica, P. O. Box 2871, Beijing 100085, P. R. China
Fax +86(10)62923563; E-mail: jning@mail.rcees.ac.cn

[b] Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100094, P. R. China

Received 12 July 2005

Abstract: An efficient method for the synthesis of 3,5-branched octaarabinoside, which is a dimer of the tetrasaccharide present in motif B of the *Mycobacterium tuberculosis* cell wall, has been developed by using 2,3,5-tri-*O*-benzoyl-α-D-arabinofuranosyl trichloroacetimidate, 5-*O*-acetyl-2,3-di-*O*-benzoyl-α-D-arabinofuranosyl trichloroacetimidate, 1,2-*O*-isopropylidene-5-*O*-trityl-β-D-arabinofuranose, dodecyl 2,3-di-*O*- benzoyl-α-D-arabinofuranoside and phenyl 2,3-di-*O*-benzoyl-1-thio-α-D-arabinofuranoside as synthons.

Key words: oligosaccharides, synthesis, glycosylations, stereoselectivity, glycosides

Saccharides have many key biological functions,[1] when conjugated to protein to form glycoproteins, they can alter protein structure and function. As components of glycolipids, they can play pivotal roles in cell–cell recognition and signaling. The extracellular matrix contains proteoglycans, large glycoconjugates that not only modify the physicochemical properties of the solution but also are involved in many recognition processes. Although numerous carbohydrate structures occur in nature, in general, the role of saccharide structure in function has been minimally studied. This can be attributed mainly to the difficulty of synthesizing saccharides. Unlike proteins and nucleic acids, saccharides are more difficult to synthesize because (i) the molecules are typically branched rather than linear, (ii) the monosaccharide unites can be connected by α- or β-linkages, and (iii) oligosaccharide synthesis requires multiple selective protection and deprotection steps. Although over the past few decades, considerable progress has been made in this field,[2] currently there is still no general route for saccharide synthesis and glycosylation chemistry is still not predicable or generally accessible. To facilitate the synthesis of target oligosaccharides, regio- and stereoselective glycosylation of glycosyl donors with unprotected or partially protected sugar acceptors has been extensively studied. Employing this strategy, in the last few years, we have prepared a lot of oligosaccharides with various structures present in natural sources such as 3,6-branched gluco-oligosaccharides,[3] 2,6-branched manno-oligosaccharides,[4] 2,6-branched,[5] 3,6-branched and 5,6-branched galacto-oligasaccharides.[6]

SYNLETT 2005, No. 15, pp 2267–2272
Advanced online publication: 07.09.2005
DOI: 10.1055/s-2005-872674; Art ID: U22205ST

Several approaches have been taken with success for the chemical synthesis of oligosaccharides.[2] Most involve the activation of the anomeric leaving group with a Lewis acid and then displacement of that leaving group by the free hydroxyl of the acceptor sugar. The Koenigs–Knor method of coupling glycosyl halides, one of the first techniques to gain widespread usage, is still in common use. Trichloroacetimidates,[7] prepared by the reaction of free sugar with trichloroacetonitrile and base, are used most frequently for coupling, e.g. glycosyl sulfoxides,[8] phosphates,[9] and phosphates[10] and thio-[11] and pentenyl glycosides.[12]

Mycobacterium (*M.*) *tuberculosis*, the causative agent of tuberculosis, belongs to the mycobacteria.[13] An impressive feature of this genus of bacteria is that they synthesize cell wall polysaccharides containing predominantly furanose residues,[13] whereas a similar phenomenon has never been found in humans. For *M. tuberculosis*, the major polysaccharides formed in its cell wall are an arabinogalactan (AG) and a lipoarabinomannan (LAM) in which all of the galactose and arabinose residues are present in the furanose form.[13] The organism's ability to make these polymers is critical to its survival, and therefore interfering with the biosynthesis of these polysaccharides is an approach for identifying anti-TB drugs.[14]

Five major motifs called A, B, C, D, and E, which are linked to the terminal ends of AG and/or LAM, have been isolated and characterized[15] as shown in Figure 1. These oligosaccharides may play an important role in the survival and pathogenicity of the organism. Since these oligosaccharides are not present in humans, they tend to be highly antigenic in the human immune system. Besides, it has been demonstrated recently that one of the anti-TB drugs used to treat tuberculosis (ethambutol) for the last 35 years actually acts by inhibiting the biosynthesis of arabinan portions of LAM and AG.[16]

Providing sufficient quantities of a sample is a basic prerequisite for detailed studies on a compound's fundamental biochemical properties and possible biological functions. However, it is very difficult to gain arabinose-containing oligosaccharides with clear structural information for biological studies from natural sources. So there has been much interest in the synthesis of the components of the *M. tuberculosis* cell wall polysaccharides. The preparations of motifs A, B, C, D, and E oligosaccharide derivatives have been finished recently.[6b,17] Compound **6**,

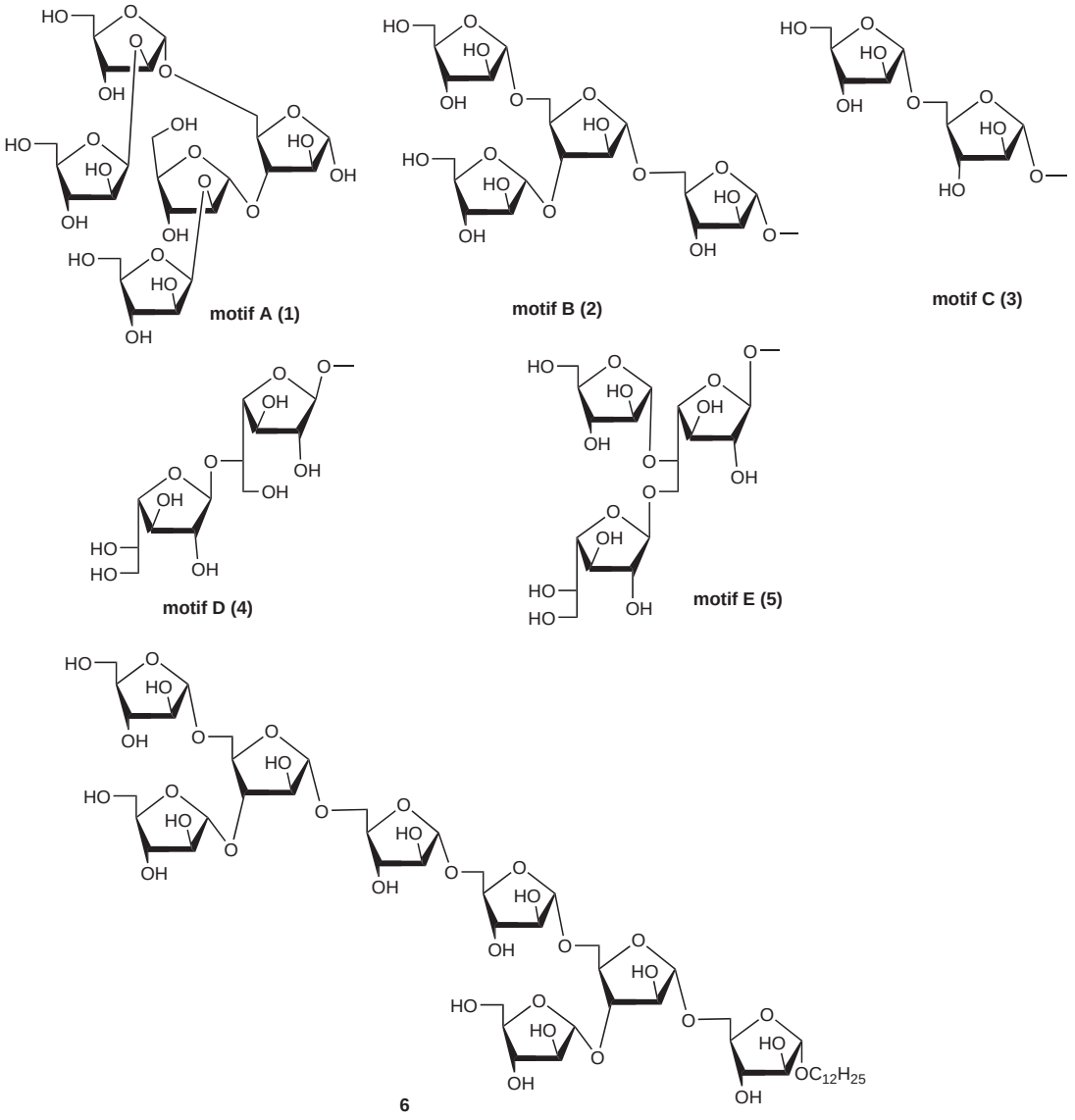


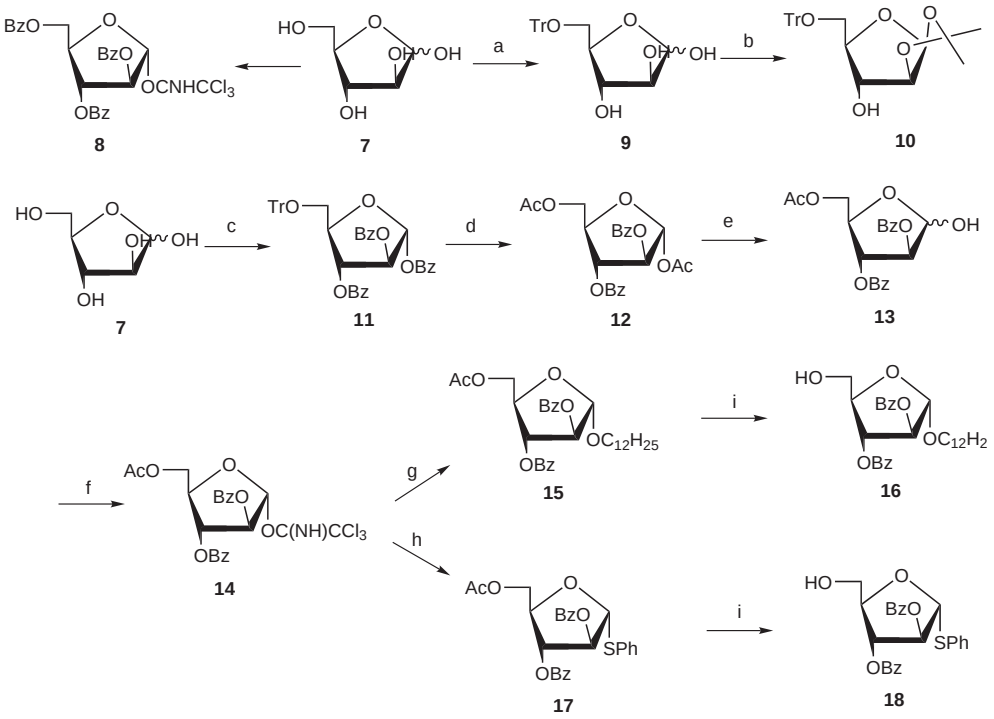
Figure 1

a dimer of the tetrasaccharide present in Motif B, is one of the arabinose-containing oligosaccharides. For detailed characterization of AG and LAM, especially for further elucidation of the molecular structure responsible for their biological properties, it would be necessary to synthesize compound **6**. This, together with the fact that the synthesis of **6** was not previously reported, prompted us to synthesize the octasaccharide using a '4+4' strategy.

To complete the synthesis of target molecule **6**, we designed a convergent strategy using tetrasaccharide **28** as acceptor and tetrasaccharide thioglycoside **26** as donor, which were assembled by the building blocks **8**, **10**, **14**, **16** and **8**, **10**, **14**, **18**, respectively.

In our synthesis, 2,3,5-tri-*O*-benzoyl-α-D-arabinofuranosyl trichloroacetimidate (**8**), 1,2-*O*-isopropylidene-5-*O*-trityl-β-D-arabinofuranose (**10**), 5-*O*-acetyl-2,3-di-*O*-benzoyl-α-D-arabinofuranosyl trichloroacetimidate (**14**), dodecyl 2,3-di-*O*-benzoyl-α-D-arabinofuranoside (**16**) and phenyl 2,3-di-*O*-benzoyl-1-thio-α-D-arabinofuranoside (**18**) were the basic building materials. Compound **8** was obtained according to the standard method.[18] Compound **10** was prepared by tritylation of D-arabinose, followed by 1,2-O-isopropylation.[19] Tritylation of D-arabinose followed by benzoylation in a one-pot manner gave 1,2,3-tri-*O*-benzoyl-5-*O*-trityl-α-D-arabinofuranose (**11**) in 71% yield for two steps (Scheme 1). Selective acetolysis of **11** using CH_2Cl_2–AcOH–Ac_2O–H_2SO_4 in a ratio of 15:10:10:1 afforded the corresponding 1,5-diacetate **12** in 83% yield. The diacetate **12** was selectively deacetylated at the anomeric position with benzylamine in THF in high yield (88%) to give the corresponding 5-*O*-acetyl-2,3-di-*O*-benzoyl-α-D-arabinofuranose (**13**).

Subsequent reaction of **13** with CCl_3CN/K_2CO_3 in CH_2Cl_2 afforded glycosyl donor **14**. Coupling of **14** with dodecyl alcohol gave compound **15** which was selectively converted to **16** using MeOH containing 0.1% HCl. The solution of MeOH containing HCl was formed in situ by



Scheme 1 *Reagents and conditions*: (a) TrCl (1.2 equiv), pyridine, 40 °C, 48 h, 64%; (b) $Me_2C(OMe)_2$, $TsOH·H_2O$, acetone, r.t., 1.5 h, 88%; (c) i. TrCl (1.2 equiv), pyridine, 40 °C, 48 h; ii. PhCOCl (3.6 equiv), r.t., 24 h, 71% (for twe steps); (d) CH_2Cl_2–HOAc–Ac_2O–H_2SO_4 = 15:10:10:1, r.t., 12 h, 83%; (e) $BnNH_2$ (4.0 equiv), THF, r.t., 24 h, 88%; (f) CH_2Cl_2, CCl_3CN (3.3 equiv), K_2CO_3 (5 equiv), r.t., 12 h, 86%; (g) dodecyl alcohol (2 equiv), TMSOTf (0.05 equiv), CH_2Cl_2, r.t., 3 h, 89%; (h) PhSH (1.5 equiv), TMSOTf (0.05 equiv), CH_2Cl_2, r.t., 2 h, 88%; (i) MeOH–HCl (0.1%), r.t., 12–14 h, 93% for **16**, 94% for **18**.

adding acetyl chloride to MeOH. Compound **18** was prepared from **14** and thiophenol in a similar fashion to that of making **16**.

Coupling of **8** with **10** followed by detritylation with solid $FeCl_3·6H_2O$ (commercial product, without any pretreatment) in CH_2Cl_2[20] gave disaccharide glycosyl acceptor **20** in 72% yield for two steps (Scheme 2). Coupling of **20** with glycosyl donor **14** using TMSOTf as catalyst and 4 Å molecular sieves in CH_2Cl_2 afforded trisaccharide **21** in 91% yield. Removal of the 1,2-*O*-isopropylidene group of **21** in 80% AcOH followed by acetylation with acetic anhydride in pyridine, selective 1-O-deacetylation with benzylamine in THF, and subsequent treatment with CCl_3CN/K_2CO_3 afforded the desired trisaccharide glycosyl donor **25** in 71% yield for four steps.

Condensation of **25** with **18** afforded tetrasaccharide thioglycoside donor **26** in 91% yield. Condensation of **25** with **16** afforded tetrasaccharide **27** in 91% yield. Selective 5-O-deacetylation of **27** in MeCOCl–MeOH–CH_2Cl_2 (1:500:500) gave tetrasaccharide acceptor **28** in 88% yield.[21]

Condensation of **26** with **28** afforded octasaccharide block **29** in 85% yield (Scheme 3). Deprotection of **29** using NH_3 in MeOH gave free octasaccharide **6** as an amorphous white solid in 90% yield.

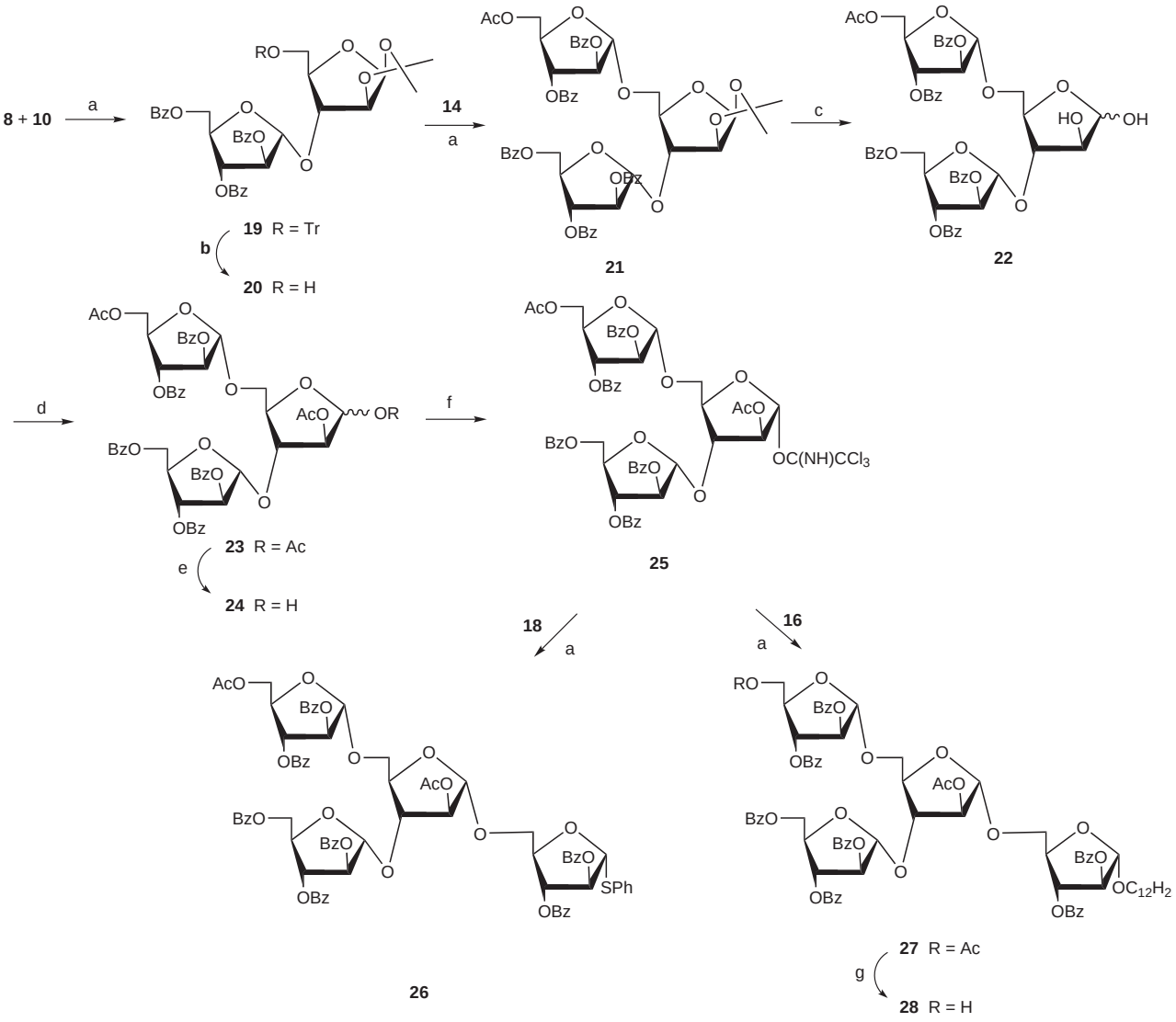
Anomeric carbons in α-arabinofuranosides resonate between δ = 105 and 110 ppm, and those of β-arabinofuranosides resonate between δ = 100 and 104 ppm.[22] Therefore, the anomeric stereochemistry of the glycosyl residues in **6** could be determined unequivocally by ^{13}C NMR spectroscopy. The ^{13}C NMR spectrum showed eight anomeric carbons at δ = 107.66, 107.61, 107.59, 107.55, 107.52, 107.43, 107.25, and 107.18 ppm, indicative of the α-linkages of all arabinofuranosyl residues in **6**.

All new compounds involved in this study were identified by optical rotations, ^1H NMR or/and ^{13}C NMR spectroscopy, and elemental analyses or MALDI-TOF MS. Selected physical data for some key compounds are given in the references.[23]

In all of the synthesis, very easily accessible materials and cheap reagents were used and the reactions were carried out smoothly in high yields and in large scales. In the synthesis, several intermediates were not separated and used directly for the further reaction simplifying the operation substantially. The sole use of acyl groups in the synthesis further simplified the procedure.

Acknowledgment

This work was supported by the National Key Project for Basic Research (2003CB114400) and by the Excellent Scientist Project from the Chinese Academy of Agricultural Sciences.



Scheme 2 *Reagents and conditions*: (a) TMSOTf (0.02 equiv), MS 4 Å, CH_2Cl_2, –42 °C, 3 h, 91%; (b) $FeCl_3 \cdot 6H_2O$ (1.5 equiv), CH_2Cl_2, r.t., 2 h, 81%; (c) 80% HOAc, reflux, 4 h, 90%; (d) Ac_2O, pyridine, r.t., 12 h, 98%; (e) $BnNH_2$ (4.0 equiv), THF, r.t., 24 h, 90%; (f) CH_2Cl_2, CCl_3CN (3.3 equiv), K_2CO_3 (5 equiv), r.t., 12 h, 90%; (g) MeCOCl–MeOH–CH_2Cl_2 (1:500:500), r.t., 12 h, 88%.

References

(1) (a) Sears, P.; Wong, C. H. *Cell. Mol. Life Sci.* **1998**, *54*, 233. (b) Varki, A. *Glycobiology* **1993**, *3*, 97. (c) Lis, H.; Sharon, N. *Eur. J. Biochem.* **1993**, *218*, 1. (d) Kolter, T.; Sandhoff, K. *Angew. Chem. Int. Ed.* **1999**, *38*, 1532.

(2) Sears, P.; Wong, C. H. *Science* **2001**, *291*, 2433.

(3) (a) Ning, J.; Yi, Y.; Kong, F. *Tetrahedron Lett.* **2002**, *43*, 5545. (b) Ning, J.; Zhang, W.; Yi, Y.; Yang, G.; Wu, Z.; Yi, Z.; Kong, F. *Bioorg. Med. Chem.* **2003**, *11*, 2193.

(4) (a) Xing, Y.; Ning, J. *Tetrahedron: Asymmetry* **2003**, *14*, 1275. (b) Ning, J.; Heng, L.; Kong, F. *Tetrahedron Lett.* **2002**, *43*, 673.

(5) Ning, J.; Yi, Y.; Yao, Z. *Synlett* **2003**, 2208.

(6) (a) Ning, J.; Wang, H.; Yi, Y. *Tetrahedron Lett.* **2002**, *43*, 7349. (b) Wang, H.; Ning, J. *J. Org. Chem.* **2003**, *68*, 2521.

(7) Rademann, J. *Angew. Chem. Int. Ed.* **1998**, *37*, 1241.

(8) Liang, R. *Science* **1996**, *274*, 1520.

(9) (a) Martin, T. J. *Tetrahedron Lett.* **1992**, *33*, 6123. (b) Kondo, H. *J. Am. Chem. Soc.* **1992**, *114*, 8748.

(10) Hashimoto, S. *J. Chem. Soc., Chem. Commun.* **1989**, 685.

(11) Ferrier, R. J. *Carbohydr. Res.* **1973**, *27*, 55.

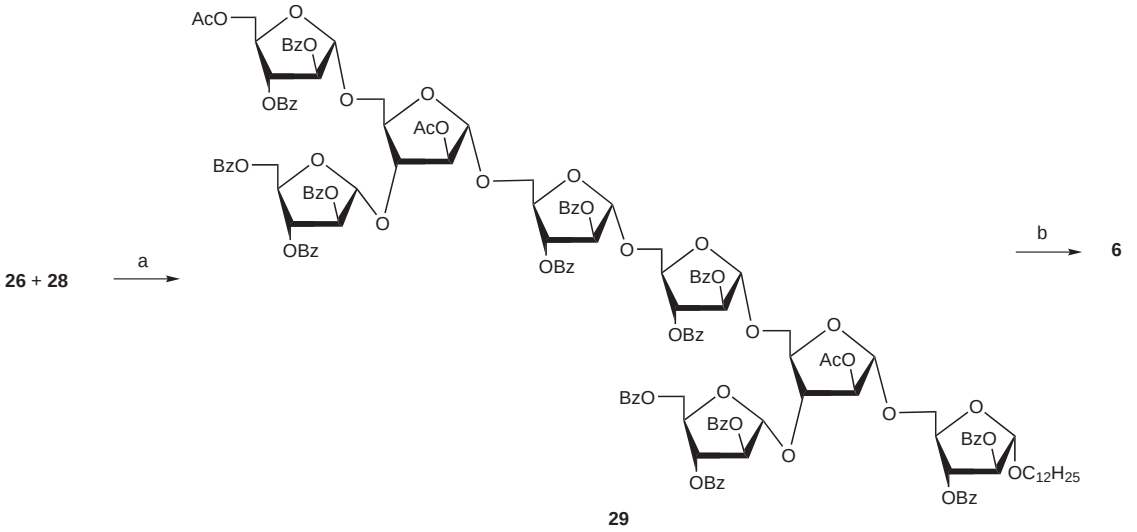
(12) Fraser-Reid, B. *Pure Appl. Chem.* **1993**, *65*, 779.

(13) Blanchard, J. S. *Annu. Rev. Biochem.* **1995**, *64*, 29.

(14) Lee, R. E.; Brennan, P. J.; Besra, G. S. *Glycobiology* **1997**, *7*, 1121.

(15) (a) Besra, G. S.; Khoo, K.-H.; McNeil, M. R.; Dell, A.; Moeeis, H. R.; Brennan, P. J. *Biochemistry* **1995**, *34*, 4257. (b) Wolika, B. A.; McNeil, M. R.; de Hoffman, E.; Chojnacki, T.; Brennan, P. J. *J. Biol. Chem.* **1994**, *269*, 23328. (c) McNeil, M. R.; Daffe, M.; Brennan, P. J. *J. Biol. Chem.* **1990**, *265*, 18200. (d) Daffe, M.; Brennan, P. J.; McNeil, M. R. *J. Biol. Chem.* **1990**, *265*, 6734. (e) McNeil, M. R.; Robuck, K. G.; Harter, M.; Brennan, P. J. *Glycobiology* **1994**, *4*, 165.

(16) (a) Lee, R. E.; Mikušová, K.; Brennan, P. J.; Besra, G. S. *J. Am. Chem. Soc.* **1995**, *117*, 11829. (b) Deng, L.; Mikušová, K.; Robuck, K. G.; Scherman, M.; Brennan, P. J.; McNeil, M. R. *Antimicrob. Agents Chemother.* **1995**, *39*, 694.



Scheme 3 *Reagents and conditions*: (a) NIS (1.5 equiv), TMSOTf (0.02 equiv), MS 4 Å, CH_2Cl_2, –42 °C, 3 h, 85%; (b) MeOH sat. with anhyd NH_3, r.t., 72 h, 90%.

(17) (a) Gurjar, M. K.; Hotha, S.; Mereyala, H. B. *Chem. Commun.* **1998**, 685. (b) D'Souza, F. W.; Lowary, T. L. *Org. Lett.* **2000**, *2*, 1493. (c) D'Souza, F. W.; Ayers, J. D.; McCarren, P. R.; Lowary, T. L. *J. Am. Chem. Soc.* **2000**, *122*, 1251. (d) Sanchez, S.; Bamhoud, T.; Prandi, J. *Tetrahedron Lett.* **2000**, *41*, 7447. (e) Gurjar, M. K.; Reddy, L. K.; Hotha, S. *Org. Lett.* **2001**, *3*, 321. (f) Gurjar, M. K.; Reddy, L. K.; Hotha, S. *J. Org. Chem.* **2001**, *66*, 4657. (g) Han, J.; Gadikota, R. R.; McCarren, P. R.; Lowary, T. L. *Carbohydr. Res.* **2003**, *338*, 581.

(18) Du, Y.; Pan, Q.; Kong, F. *Carbohydr. Res.* **2000**, *323*, 28.

(19) Izumi, M.; Tsuruta, O.; Hashimoto, H. *Carbohydr. Res.* **1996**, *280*, 287.

(20) Ding, X.; Wang, W.; Kong, F. *Carbohydr. Res.* **1997**, *303*, 445.

(21) Zeng, Y.; Li, A.; Kong, F. *Tetrahedron Lett.* **2003**, *44*, 8325.

(22) Mizutani, K.; Kasai, R.; Nakamura, M.; Tanaka, O.; Matsuura, H. *Carbohydr. Res.* **1989**, *185*, 27.

(23) All new compounds gave satisfactory elemental analysis results. Selected physical data for some key compounds are as follows:

For **14**: $[\alpha]_D$ +80.9 (*c* 1.0, $CHCl_3$). ^1H NMR (400 MHz, $CDCl_3$): δ = 8.73 (s, 1 H, CN*H*CCl_3), 8.13–7.26 (m, 10 H, 2 × Ph*H*), 6.64 (s, 1 H, H-1), 5.78 (d, 1 H, $J_{2,3}$ = 1.1 Hz, H-2), 5.51 (dd, 1 H, $J_{2,3}$ = 1.1 Hz, $J_{3,4}$ = 3.3 Hz, H-3), 4.68 (m, 1 H, H-4), 4.58 (dd, 1 H, $J_{4,5a}$ = 4.2 Hz, $J_{5a,5b}$ = 11.9 Hz, H-5a), 4.44 (dd, 1 H, $J_{4,5b}$ = 5.51 Hz, $J_{5a,5b}$ = 11.9 Hz, H-5b), 2.04 (s, 3 H, C*H*$_3$CO). Anal. Calcd for $C_{24}H_{22}Cl_3NO_8$: C, 51.59; H, 3.97. Found: C, 51.74; H, 3.86.

For **16**: $[\alpha]_D$ +98.1 (*c* 1.0, $CHCl_3$). ^1H NMR (400 MHz, $CDCl_3$): δ = 8.09–7.26 (m, 10 H, 2 × Ph*H*), 5.52 (d, 1 H, $J_{2,3}$ = 1.2 Hz, H-2), 5.42 (m, 1 H, H-3), 5.24 (s, 1 H, H-1), 4.31 (m, 1 H, H-4), 4.02–3.97 (m, 2 H, C*H*$_2$C$_{11}$H$_{23}$), 3.75 (m, 1 H, 5a), 3.53 (m, 1 H, 5b), 1.65–0.82 (m, 23 H, $CH_2C_{11}H_{23}$). Anal. Calcd for $C_{32}H_{44}O_7$: C, 71.08; H, 8.20. Found: C, 71.44; H, 8.06.

For **18**: $[\alpha]_D$ +101.1 (*c* 1.0, $CHCl_3$). ^1H NMR (400 MHz, $CDCl_3$): δ = 8.14–7.26 (m, 15 H, 3 × Ph*H*), 5.80 (s, 1 H, H-1), 5.74 (d, 1 H, $J_{2,3}$ = 1.4 Hz, H-2), 5.55 (d, 1 H, $J_{3,4}$ = 4.6 Hz, H-3), 4.59 (m, 1 H, H-4), 4.04 (m, 2 H, H-5a,b). Anal. Calcd for $C_{26}H_{24}O_6S$: C, 67.22; H, 5.21. Found: C, 67.04; H, 5.06.

For **20**: $[\alpha]_D$ +44.7 (*c* 1.0, $CHCl_3$). ^1H NMR (400 MHz, $CDCl_3$): δ = 8.06–7.26 (m, 15 H, 3 × Ph*H*), 5.96 (d, 1 H, $J_{3,4}$ = 4.1 Hz, H-3), 5.61 (d, 1 H, *J* = 4.6 Hz, H-2), 5.50 (d, 1 H, *J* = 1.2 Hz, H-1), 5.45 (s, 1 H, H-1), 4.81 (dd, 1 H, *J* = 3.7, 11.9 Hz, H-5), 4.74 (d, 1 H, *J* = 3.9 Hz, H-3), 4.68 (dd, 1 H, *J* = 4.7, 11.9 Hz, H-5), 4.58 (m, 1 H, H-5), 4.35 (d, 1 H, *J* = 3.2 Hz, H-2), 4.24 (m, 1 H, H-4), 3.80 (m, 2 H, H-4, H-5), 1.53, 1.34 (2 s, 6 H, 2 C*H*$_3$). Anal. Calcd for $C_{35}H_{38}O_{12}$: C, 64.61; H, 5.89. Found: C, 64.50; H, 5.74.

For **21**: $[\alpha]_D$ +21.9 (*c* 1.0, $CHCl_3$). ^1H NMR (400 MHz, $CDCl_3$): δ = 8.06–7.26 (m, 25 H, 5 × Ph*H*), 5.94 (d, 1 H, H-1), 5.59 (d, 1 H, *J* = 4.8 Hz, H-3), 5.50 (d, 1 H, *J* = 1.1 Hz, H-2), 5.45 (d, 1 H, *J* = 1.2 Hz, H-2), 5.38 (d, 1 H, *J* = 4.1 Hz, H-3), 5.31 (s, 1 H, H-1), 5.27 (s, 1 H, H-1), 4.72–4.51 (m, 6 H, H-2, H-4, 4 × H-5), 4.46 (d, 1 H, *J* = 3.5 Hz, H-3), 4.36 (m, 1 H, H-4), 4.28 (m, 1 H, H-4), 4.04 (dd, 1 H, *J* = 5.4, 10.4 Hz, H-5), 3.75 (dd, 1 H, *J* = 4.6, 10.4 Hz, H-5), 2.02 (s, 3 H, C*H*$_3$CO), 1.56, 1.34 [2 s, 6 H, C(C*H*$_3$)$_2$]. Anal. Calcd for $C_{57}H_{58}O_{19}$: C, 65.38; H, 5.58. Found: C, 65.49; H, 5.41.

For **25**: $[\alpha]_D$ +34.4 (*c* 1.0, $CHCl_3$). ^1H NMR (400 MHz, $CDCl_3$): δ = 8.56 (s, 1 H, CN*H*CCl_3), 8.04–7.26 (m, 25 H, 5 × Ph*H*), 6.37 (s, 1 H, H-1), 5.58 (d, 1 H, *J* = 4.8 Hz, H-3), 5.54 (d, 1 H, *J* = 1.1 Hz, H-2), 5.49 (d, 1 H, *J* = 1.2 Hz, H-2), 5.37 (d, 1 H, *J* = 4.1 Hz, H-3), 5.30 (s, 1 H, H-1), 5.28 (m, 1 H, H-2), 5.24 (s, 1 H, H-1), 4.76–4.50 (m, 6 H, H-3, H-4, 4 × H-5), 4.34 (m, 1 H, H-4), 4.11 (m, 1 H, H-4), 3.90 (dd, 1 H, *J* = 5.1, 10.4 Hz, H-5), 3.80 (dd, 1 H, *J* = 4.8, 10.4 Hz, H-5), 2.02, 2.00 (2 s, 6 H, 2 × C*H*$_3$CO). Anal. Calcd for $C_{59}H_{56}Cl_3NO_{20}$: C, 58.79; H, 4.68. Found: C, 59.88; H, 5.24.

For **26**: $[\alpha]_D$ +20.1 (*c* 1.0, $CHCl_3$). ^1H NMR (400 MHz, $CDCl_3$): δ = 8.04–7.26 (m, 40 H, 8 × Ph*H*), 5.80 (s, 1 H, H-1), 5.72 (d, 1 H, *J* = 4.1 Hz, H-3), 5.61 (m, 1 H, H-2), 5.53 (d, 1 H, *J* = 1.1 Hz, H-2), 5.47–5.45 (m, 3 H, H-1, H-2, H-3), 5.37 (s, 1 H, H-1), 5.34 (d, 1 H, *J* = 3.7 Hz, H-3), 5.27 (d, 1 H, *J* = 1.2 Hz, H-2), 5.21 (s, 1 H, H-1), 4.74 (dd, 1 H, *J* = 3.5, 11.9 Hz, H-5), 4.65–4.33 (m, 8 H, H-3, 4 H-4, 3 × H-5), 4.18 (dd, 1 H, *J* = 4.0, 11.9 Hz, H-5), 4.05 (dd, 1 H, *J* = 4.1, 11.9 Hz, H-5), 3.88 (m, 2 H, 2 × H-5), 2.01, 1.97 (2 × s, 6 H, 2 × C*H*$_3$CO). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 170.60, 169.94 (2 × CH_3*C*O), 165.96, 165.63, 165.57, 165.35, 165.31, 164.96, 164.91 (7 × Ph*C*O), 105.97, 105.69, 105.46, 91.09 (4 × C-1), 20.68, 20.67 (2 × *C*H$_3$CO). Anal. Calcd for

$C_{83}H_{78}O_{25}S$: C, 66.13; H, 5.22. Found: C, 66.41; H, 5.04.
For **27**: $[\alpha]_D$ +69.4 (*c* 1.0, $CHCl_3$). ^1H NMR (400 MHz, $CDCl_3$): δ = 8.03–7.22 (m, 35 H, 7 × Ph*H*), 5.59 (d, 1 H, *J* = 4.6 Hz, H-3), 5.53 (d, 1 H, H-2), 5.48 (s, 1 H, H-1), 5.46 (d, 1 H, *J* = 4.6 Hz, H-3), 5.44 (d, 1 H, *J* = 1.2 Hz, H-2), 5.41 (d, 1 H, *J* = 1.1 Hz, H-2), 5.37 (s, 1 H, H-1), 5.34 (d, 1 H, *J* = 3.4 Hz, H-3), 5.28 (d, 1 H, *J* = 1.3 Hz, H-2), 5.21 (d, 2 H, *J* = 4.7 Hz, 2 × H-1), 4.73 (dd, 1 H, *J* = 3.6, 11.6 Hz, H-5), 4.63 (m, 1 H, H-5), 4.53–4.33 (m, 7 H, H-3, 2 × H-4, 2 × H-5, $CH_2C_{11}H_{23}$), 4.13 (m, 1 H, H-5), 4.05 (dd, 1 H, *J* = 3.9, 11.6 Hz, H-5), 3.91 (dd, 1 H, *J* = 3.1, 11.6 Hz, H-5), 3.83 (dd, 1 H, *J* = 2.8, 11.9 Hz, H-5), 3.74 (m, 1 H, H-4), 3.49 (m,1 H, H-4), 2.00, 1.98 (2 × s, 6 H, 2 × CH_3CO), 1.61–0.82 (m, 23 H, $CH_2C_{11}H_{23}$). Anal. Calcd for $C_{89}H_{98}O_{26}$: C, 67.50; H, 6.24. Found: C, 67.66; H, 6.09.
For **28**: $[\alpha]_D$ +74.4 (*c* 1.0, $CHCl_3$). ^1H NMR (400 MHz, $CDCl_3$): δ = 8.04–7.26 (m, 35 H, 7 × Ph*H*), 5.60 (d, 1 H, *J* = 4.6 Hz, H-3), 5.57 (d, 1 H, *J* = 1.3 Hz, H-2), 5.4 (m, 2 H, H-1, H-3), 5.43 (d, 1 H, *J* = 1.2 Hz, H-2), 5.40 (d, 1 H, *J* = 1.1 Hz, H-2), 5.35 (m, 2 H, H-1, H-3), 5.29 (d, 1 H, 1.4 Hz, H-2), 5.22 (s, 1 H, H-1), 5.21 (s, 1 H, H-1), 4.78 (dd, 1 H, *J* = 3.6, 11.9 Hz, H-5), 4.65 (dd, 1 H, *J* = 4.6, 11.9 Hz, H-5), 4.59 (m, 1 H, H-4), 4.50 (m, 1 H, H-4), 4.38–4.32 (m, 3 H, H-3, 2 × H-5), 4.16 (dd, 1 H, *J* = 4.3, 11.2 Hz, H-5), 3.99–3.92 (m, 4 H, 2 × H-5, $CH_2C_{11}H_{23}$), 3.84 (dd, 1 H, *J* = 3.5, 11.2 Hz, H-5), 3.74 (m, 1 H, H-4), 3.50 (m,1 H, H-4), 2.00 (s, H, CH_3CO), 1.60–0.86 (m, 23 H, $CH_2C_{11}H_{23}$). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 169.92 (CH_3CO), 166.16, 165.90, 165.67, 165.54, 165.42, 165.01, 164.86 (7 × Ph*C*O), 105.74, 105.46, 105.41, 105.42 (4 × C-1), 20.67 (CH_3CO). Anal. Calcd for $C_{87}H_{96}O_{25}$: C, 67.78; H, 6.28. Found: C, 67.60; H, 6.41.
For **29**: $[\alpha]_D$ +69.4 (*c* 1.0, $CHCl_3$). ^1H NMR (400 MHz, $CDCl_3$): δ = 8.01–7.26 (m, 70 H, 14 × Ph*H*), 5.51 (s, H-1), 5.50 (s, H-1), 5.45–5.40 (m, 3 H, H-1), 5.33 (s, H-1), 5.45–5.40 (m, 2 H, H-1), 1.99, 1.97, 1.95 (3 × s, 9 H, 3 × CH_3CO), 1.57–0.85 (m, 23 H, $CH_2C_{11}H_{23}$). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 170.58, 169.98, 169.91 (3 × CH_3CO), 165.93, 169.94, 165.62, 165.63, 165.55, 165.51, 165.49, 165.41, 165.42, 165.12, 165.04, 164.93, 164.88, 164.84 (14 × Ph*C*O), 106.03, 105.92, 105.79, 105.61, 105.59, 105.52, 105.47, 105.42 (8 × C-1), 20.70, 20.69, 20.66 (3 × CH_3CO). Anal. Calcd for $C_{164}H_{168}O_{50}$: C, 67.02; H, 5.76. Found: C, 67.11; H, 5.60.
For **6**: $[\alpha]_D$ –18.9 (*c* 1.0, H_2O). ^1H NMR (400 MHz, D_2O): δ = 5.16 (s, 2 H, H-1), 5.10 (s, 5 H, H-1), 4.96 (s, 1 H, H-1). 1.62–0.90 (m, 23 H, $CH_2C_{11}H_{23}$). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 107.66, 107.61, 107.59, 107.55, 107.52, 107.43, 107.25, 107.18 (8 × C-1). MALDI-TOF MS: *m/z* calcd for $C_{60}H_{106}O_{33}$: 1355.46 [M], found: 1378.09 $[M + Na]^+$.