ORGANIC LETTERS

2001
Vol. 3, No. 21
3265−3268

Chemoenzymatic Iterative Synthesis of Difficult Linkages of Oligosaccharides on Soluble Polymeric Supports

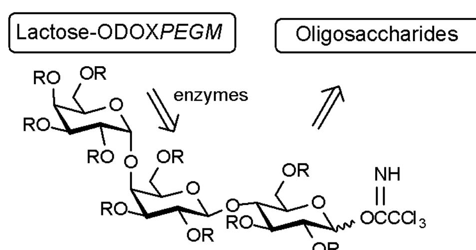
Fengyang Yan, Michel Gilbert, Warren W. Wakarchuk, Jean-Robert Brisson, and Dennis M. Whitfield*

Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Room 3024, Ottawa, Ontario K1A 0R6, Canada

dennis.whitfield@nrc.ca

Received July 20, 2001

ABSTRACT



A trisaccharide donor containing a *cis*-Gal*p*α(1→4)Gal*p* linkage was prepared using a synthetic strategy based on chemoenzymatic oligosaccharide synthesis on a soluble polymeric support. Significantly, only retaining glycosyltransferases gave complete reactions, whereas inverting enzymes showed little or no activity with poly(ethylene glycol) (MPEG)-bound lactose as an acceptor. The MPEG-attached trisaccharide was shown to bind to Verotoxin-1 by transfer NOE studies through the Gal*p*α(1→4)Gal*p* portion of the molecule.

Our research activities are directed toward the simplification of the synthesis of oligosaccharides. Toward this goal, we have developed methodological improvements to polymer-supported and chemoenzymatic syntheses. One of the major unsolved problems in polymer-supported oligosaccharide synthesis is how to incorporate "difficult" linkages such as cis linkages.[1] One strategy is to prepare oligosaccharide building blocks containing the "difficult" linkage and use the building block in the preparation of larger oligosaccharides. We demonstrate that oligosaccharide building blocks containing cis linkages can be synthesized on polymeric supports chemoenzymatically. This strategy combines the advantage of the absolute stereochemical control of *cis*-glycoside bond formation afforded by glycosyltransferases[2] and the simple purifications from using polymeric supports.[3] These efforts culminated in the synthesis of the P^k antigen as a trisaccharide and a tetrasaccharide attached to the polymer poly(ethylene glycol) (MPEG), $CH_3O(CH_2CH_2O)_nH$ via the linker dioxyxylene [DOX, −(O)CH_2PhCH_2(O)−].[4] These oligosaccharides contain a *cis*-Gal*p*α(1→4)Gal*p* linkage and are of interest because they could bind to the Verotoxin from *Escherichia coli* O157.[5]

Initially, lactose was glycosidically bound to (MPEG)-(DOX)OH **1** by glycosylation with the perbenzoylated

* To whom correspondence should be addressed. Phone: (613) 993-5265. Fax: (613) 952-9092.

(1) (a) Belogi, G.; Zhu, T.; Boons, G.-J. *Tetrahedron Lett.* **2000**, *41*, 6969. (b) Flitsch, S. L.; Watt, G. M. *Glycopeptides and Related Compounds*; Large, D. G., Warren, C. D., Ed.; Marcel Dekker: New York, 1997; p 207.

(2) (a) Endo, T.; Koizumi, S. *Curr. Opin. Struct. Biol.* **2000**, *10*, 536. (b) Windmuller, R.; Schmidt, R. R. *Tetrahedron Lett.* **1994**, *35*, 7927.

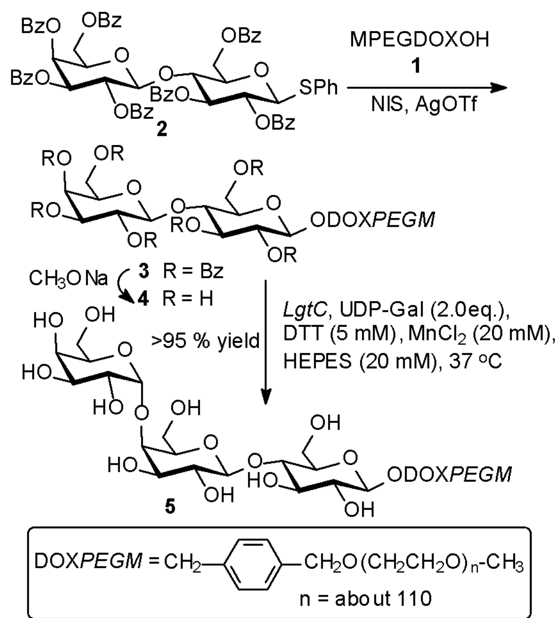
(3) (a)Krepinsky, J. J.; Douglas, S. P. *Carbohydrates in Chemistry and Biology, Part I: Chemistry of Saccharides*; Ernst, B., Hart, G. W., Sinay, P., Ed.; Wiley-VCH: Weinheim, 2000. (b) Haase, W.-C.; Seeberger, P. H. *Curr. Org. Chem.* **2000**, *4*, 481.

(4) Douglas, S. P.; Whitfield, D. M.; Krepinsky, J. J. *J. Am. Chem. Soc.* **1995**, *117*, 2116.

(5) (a) Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature* **2000**, *403*, 669. (b) Fan, E.; Merritt, E. A.; Verlinde, C. L. M. J.; Hol, W. G. J. *Curr. Opin. Struct. Biol.* **2000**, *10*, 680.

10.1021/ol016466j CCC: $20.00 © 2001 American Chemical Society
Published on Web 09/18/2001

Scheme 1



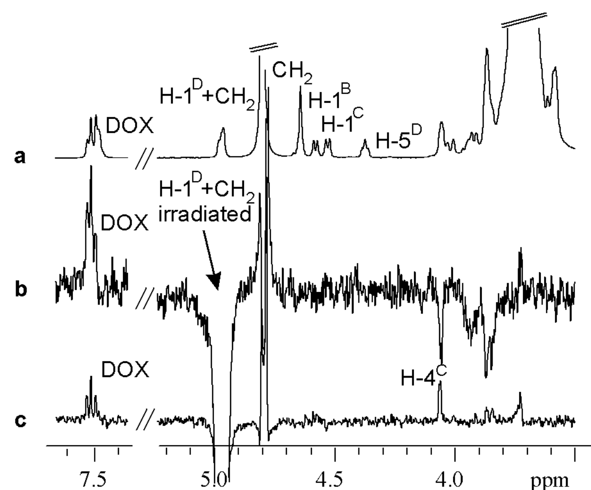
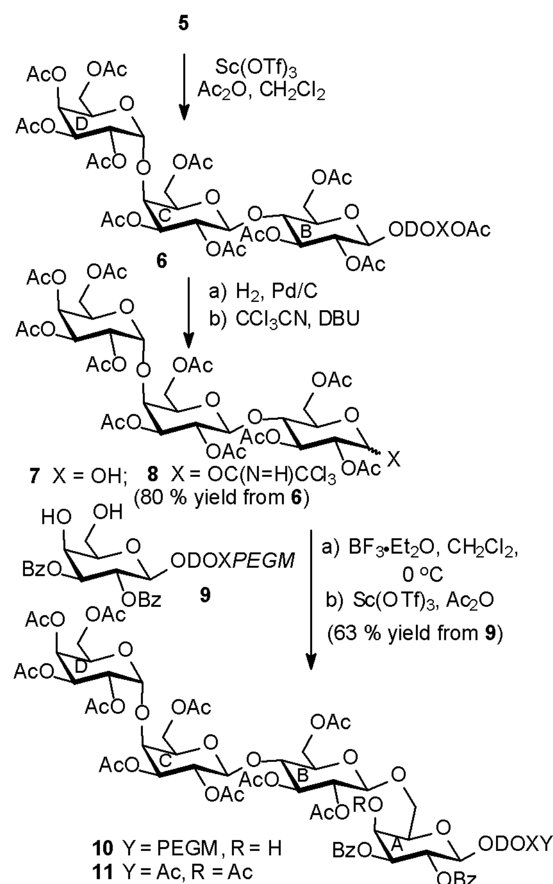


Figure 1. Partial one-dimensional transient NOE spectra of 2 mM polymer-bound trisaccharide **5** and **5** with Verotoxin-1 (200/1) in D_2O: (a) ^1H NMR spectra of **5** in D_2O; (b) 1D NOESY spectra of **5** and VT-1 in D_2O; (c) 1D NOESY spectra of **5** in D_2O.

thioglycoside donor **2** (see Scheme 1). All attempts to use the more accessible peracetylated lactose donors were frustrated by concomitant acetyl transfer to the polymer-bound construct.[6] Reactions with **2** proceed in essentially quantitative yield to **3** free of this side reaction. The soluble polymer-bound compounds can be purified by precipitation with *tert*-butyl methyl ether (TBME) and reprecipitation from absolute ethanol. After debenzoylation, the acceptor **4** can be dissolved in aqueous buffers since (MPEG) is also readily water soluble. Thus, the well-characterized enzyme α(1→4)-galactosyltransferase (*LgtC*)[7] and the nucleotide donor UDP-Gal can be incubated with **4**. The reaction proceeds to near completion to yield polymer-bound P^k trisaccharide **5**. Purification involves centrifugation to remove insoluble precipitates, lyophilization followed by extraction into dichloromethane, filtration, and evaporation. This material is nearly pure and can be further purified by reprecipitation from ethanol. In some cases, treatment with acidic ion-exchange resin is recommended to remove traces of paramagnetic metals (e.g., Mn^{2+}), which can interfere with NMR studies. NMR studies are conveniently performed in $CDCl_3$ or D_2O solutions. Signals for **4** were not detectable, and only those consistent with **5** were observed (see Figure 1a).

The trisaccharide can be cleaved from the (MPEG)(DOX) by $Sc(OTf)_3$ in the presence of acetic anhydride to yield easily purifiable peracetylated sugar−DOXOAc products. These procedures are easily scaled up to yield hundreds of milligrams of trisaccharide **6** (see Scheme 2).[8] A sequence of hydrogenation to yield **7** followed by treatment with CCl_3-CN and DBU leads to the trichloroacetimidyl donor **8**. This donor can be used to glycosylate polymer bound acceptor **9** to yield **10** and after cleavage to yield tetrasaccharide **11**.[9] Thus, the *cis*-Gal*p*α(1→4)Gal*p* linkage of **5** is incorporated into a polymer-bound oligosaccharide. This sequence is a

Scheme 2



(6) Nukada, T.; Berces, A.; Whitfield, D. M. *J. Org. Chem.* **1999**, *64*, 9030.
(7) Wakarchuk, W. W.; Cunningham, A.; Watson D. C.; Young, N. M. *Protein Eng.* **1998**, *11*, 295.

prototype of a powerful strategy to overcome one of the inherent problems in synthesizing oligosaccharides by polymer-supported methods, namely absolute stereochemical control of *cis*-glycoside formation. This new iterative protocol extends our previous approaches to this problem based on chemoenzymatic synthesis of oligosaccharide donors[10] and complements a recent entirely chemical approach, which relies on developing a 100% stereoselective glycosylation reaction.[11]

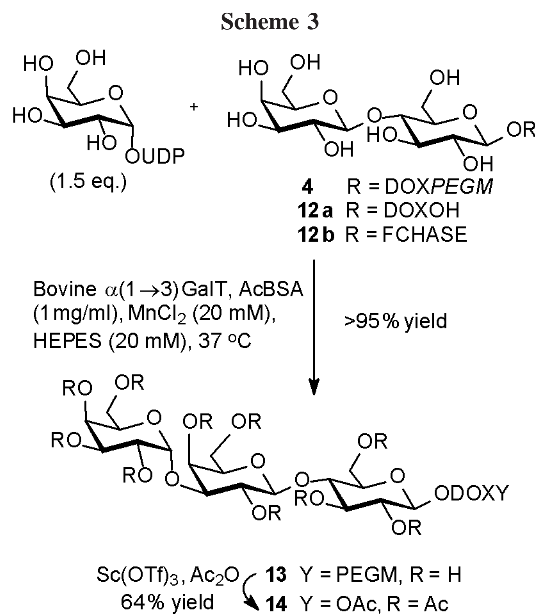
This chemoenzymatic protocol relies on finding enzymes that are compatible with MPEG-bound acceptors. We have examined a number of glycosyltransferases, which include so-called retaining and inverting glycosyltransferases, for this activity (see Table 1). Experiments with the inverting enzymes, α(2→3)sialyltransferases and β(1→3)-*N*-acetylglucosaminyltransferase (*LgtA*) from *Neisseria meningitidis* or *Campylobacter jejuni* gave extremely low yields or no reaction.[12] Typically, the reactions were treated with more enzyme and excesses of nucleotide sugar donors. However, the glycosylation reaction yields were not increased. *LgtC* is a retaining enzyme, so we tested **4** as a substrate for another retaining enzyme, the bovine α(1→3)galactosyl transferase.[13] As expected, near-quantitative galactosylation of the MPEG-bound acceptor **4** was achieved in the presence of only a small excess of uridine 5′-diphospho galactose (UDP-Gal, 1.5 equiv). The 1D and 2D NMR spectroscopy (gCOSY, gHSQC, and HMBC) of **14** indicated that the glycosidic linkage is a *cis*-Gal*p*α(1→3)Gal*p* linkage (chemical shift of terminal Gal anomeric proton: δ = 5.23 ppm, *J* = 2.0 Hz). This allows for the synthesis of the so-called xenotransplantation antigen Gal*p*α(1→3)Gal*p*[14] (see Scheme 3).

Table 1. Glycosylation Reaction of Gal*p*α(1→4)Glc*p*β-*O*-(DOX)(PEGM) **4** with a Number of Retaining and Inverting Glycosyltransferases

enzyme (gene)	donor (α/β)	yield (%) (config (α/β))
α(2→3) sialyltransferase	CMP-NeuNAc (β)	~0
(NST−27)[a]		inversion (α)
(CST−04)[b]		~0−2
		inversion (α)
(CST−06)[b]		~0−5
		inversion (α)
β(1→3)-*N*-acetylglucosaminyl transferase (LgtA)[c]	UPD-GlcNAc (α)	~0−5 inversion (β)
α(1→3)galactosyltransferase (bovine)	UDP-Gal (α)	>95 retention (α)
α(1→4)galactosyltransferase (LgtC)		>95 retention (α)

[a] For reaction conditions, see ref 20. [b] Acceptor **4** (1 mM), HEPES (50 mM), MgCl$_2$ (20 mM), CMP-Neu5Ac (2 mM), 0.08 U of enzyme, 37 °C, 18 h. [c] For reaction conditions, see ref 20.

Scheme 3



A recent report successfully used glycosidases with our (MPEG)(DOX) system, but the maximum reported yields were 24%, which would compromise our new iterative

(8) The stereo- and regiochemistry of α(1→4) linkage **6** was confirmed by 1D and 2D NMR spectroscopy (gCOSY, gHSQC, and HMBC). Selected ^{1}H and ^{13}C NMR data for compound **6**: ^{1}H NMR (500 MHz, CDCl$_3$) δ 3.58−3.64 (m, 1H, H-5^B), 3.75 (t, 1H *J* = 7.0 Hz, H-5^C), 3.82 (t, 1H, *J* = 9.0 Hz, H-4^B), 4.00 (d, 1H, *J* = 2.0 Hz, H-4^C), 4.08−4.18 (m, 4H, H-6^B, H-6^C, 2 × H-6^D), 4.46−4.51 (m, 2H, H-5^D, H-6^B), 4.51 (d, 1H, *J* = 8.0 Hz, H-1^C), 4.52 (d, 1H *J* = 8.0 Hz, H-1^B), 4.60 (d, 1H, *J* = 12.5 Hz, CHDOX), 4.73 (dd, 1H, *J* = 3.0, 10.5 Hz, H-3^C), 4.86 (d, 1H, *J* = 12.0 Hz, CHDOX), 4.96 (dd, 1H, *J* = 8.5, 10.0 Hz, H-2^B), 4.98 (d, 1H, *J* = 3.5 Hz, H-1^D), 5.07 (dd, 1H, *J* = 8.0, 9.2 Hz, H-2^C), 5.09 (s, 2H, DOXCH$_2$), 5.16 (t, 1H, *J* = 9.0 Hz, H-3^B), 5.17 (dd, 1H, *J* = 3.5, 11.0 Hz, H-2^D), 5.38 (dd, 1H, *J* = 3.5, 11.0 Hz, H-3^D), 5.58 (d, 1H, *J* = 2.5 Hz, H-4^D); ^{13}C NMR (50.32 MHz, CDCl$_3$,) δ 60.24 (C-6^D), 61.26 (C-6^C), 62.12 (C-6^B), 65.90 (CH$_2$), 67.04 (C-5^D), 67.10 (C-3^D), 67.86 (C-4^D), 68.80 (C-2^D), 68.95 (C-2^C), 70.26 (CH$_2$), 71.71 (C-5^C), 71.78 (C-2^B), 72.57 (C-5^B), 72.76 (C-3^C), 73.08 (C-3^B), 76.37 (C-4^B), 76.88 (C-4^C), 99.00 (C-1^B), 99.59 (C-1^D), 101.03 (C-1^C).

(9) The β(1→6) regio- and stereochemistry of **11** was confirmed by measuring 1D and 2D NMR spectroscopy (gCOSY, gHSQC, and HMBC). Selected ^{1}H NMR and MS data for compound **11**: ^{1}H and ^{13}C NMR (500 MHz, CDCl$_3$) δ 3.62−3.66 (m, 1H, H-5^B), 3.73−3.83 (m, 3H, H-5^C, H-6^A, H-4^B), 3.89 (dd, 1H, *J* = 4.0, 10.5 Hz, H-6^A), 3.96−4.05 (m, 2H, H-5^A, H-4^C), 4.09−4.19 (m, 4H, H-6^B, 2 × H-6^D, H-6^C), 4.41−4.46 (m, 1H, H-6^C), 4.46−4.54 (m, 2H, H-6^B, H-5^D), 4.52 (d, 1H, *J* = 7.5 Hz, H-1^C), 4.59 (d, 1H, *J* = 8.0 Hz, H-1^B), 4.69 (d, 1H, *J* = 8.5 Hz, H-1^A), 4.69 (d, 1H, *J* = 11.5 Hz, CHDOX), 4.73 (dd, 1H, *J* = 2.0, 10.5 Hz, H-3^C), 4.92 (dd, 1H, *J* = 8.5, 10.0 Hz, H-2^B), 4.93 (d, 1H, *J* = 12.0 Hz, CHDOX), 4.98 (1H, *J* = 3.0 Hz, H-1^D), 5.04 (s, 2H, DOXCH$_2$), 5.11 (dd, 1H, *J* = 7.5, 10.5 Hz, H-2^C), 5.18 (dd, 1H, *J* = 3.5, 11.0 Hz, H-2^D), 5.20 (t, 1H, *J* = 9.5 Hz, H-3^B), 5.33 (dd, 1H, *J* = 3.0, 10.5 Hz, H-3^A), 5.39 (dd, 1H, *J* = 3.0, 10.5 Hz, H-3^D), 5.59 (2H, H-4^A, H-4^D), 5.72 (dd, 1H, *J* = 8.0, 10.0 Hz, H-2^A); ^{13}C NMR (50.32 MHz, CDCl$_3$) δ 60.29(C-6^D), 61.37 (C-6^C), 62.09 (C-6^B), 65.96 (CH$_2$), 67.09 (C-3^D), 67.17 (C-5^D), 67.90 (C-6^A), 67.89 (C-4^D), 68.98 (C-4^A), 69.52 (C-2^D), 69.62 (C-2^C), 69.95 (C-2^A), 70.00 (CH$_2$), 71.66 (C-2^B), 71.79 (C-3^A), 71.88 (C-5^C), 72.73 (C-5^B), 72.74 (C-5^A), 72.85 (C-3^C), 73.05 (C-3^B), 76.32 (C-4^B), 77.23 (C-4^C), 99.48 (C-1^A), 99.67 (C-1^D), 100.27 (C-1^B), 101.03 (C-1^C); MS (MALDI) calcd for $C_{70}H_{82}O_{36}Na$ 1521.45, found *m/z* 1521.12 (M + Na$^+$).

(10) Mehta, S.; Gilbert, M.; Wakarchuk, W. W.; Whitfield, D. M. *Org. Lett.* **2000**, *2*, 751.

(11) Zhu, T.; Boons, G.-J. *J. Am. Chem. Soc.* **2000**, *122*, 10222.

(12) (a) Wakarchuk, W. W.; Martin, A.; Jennings, M. P.; Moxon, E. R.; Richards, J. C. *J. Biol. Chem.* **1996**, *271*, 19166, (b) α(2→3)Sialyltransferases from *C. jejuni*: Gilbert, M.; Brisson, J.-R.; Karwaski, M.-F.; Michniewicz, J.; Cunningham, A.-M.; Wu, Y.; Young, N. M.; Wakarchuk, W. W. *J. Biol. Chem.* **2000**, *275*, 3896.

(13) Blanken, W. M.; Van den Eijnden, D. H. *J. Biol. Chem.* **1985**, *260*, 12927.

protocol.[15] A recent report using a dendritic linker system attached to a PEG derivative also encountered difficulties in obtaining quantitative yields.[16] The origin of this inhibition of the transferases by MPEG-bound acceptors is not clear. We have done a number of control experiments that rule out simple explanations. For example, the deacetylated lactoseDOX glycoside **12a** is an acceptor for some of the inactive transferases, ruling out a simple conformational explanation. Also, incubating a standard lactose-FCHASE[17] acceptor **12b** in the presence of (MPEG)(DOX)OH **1** does not inhibit the inverting transferase of α(2→3)sialyltransferase from *N. meningitidis*. Similarly, incubating **12a** in the presence of **4** and the α(2→3)sialyltransferase from *C. jejuni* (CST-06) did not inhibit the sialylation of **12a**, while **4** was not sialylated. These direct competition experiments rule out nonspecific inhibition by PEG.

We currently hypothesize a "strangulation" effect in which binding of the acceptor to the enzyme active site leads to metastable binding of the PEG chains to the protein and hence inhibition of transferase activity. Without the nucleation by sugar–protein binding, the PEG–protein interaction is too weak to lead to enzyme inhibition. We are currently examining more transferases and testing different experimental protocols to optimize these reactions.

To study the mechanism(s) of MPEG-attached oligosaccharides binding to proteins, we performed NMR studies of **5** in D_2O solution. Construct **5** was shown to bind to the Verotoxin-1 from *E. coli* O157[18] by ^{1}H NMR transfer NOE measurements.[19] In the absence of protein, **5** behaves like a small molecule with $\omega\tau_c < 1$ and, hence, positive NOEs; see Figure 1c. Once bound to the protein, the NOEs across the Gal*p*α(1→4)Gal*p* linkage become negative ($\omega\tau_c > 1$), whereas those in the DOX linker remain positive; see Figure 1b. This clearly indicates that the major binding is between the sugar and the protein and not the linker or the PEG. Furthermore, without sophisticated numerical analyses of the NOE data, one can conclude that the Gal*p*α(1→4)Gal*p* portion of the molecule binds to the protein in agreement with previous NMR studies.[19] These interesting motional properties of MPEG-bound oligosaccharides are currently being studied in more detail to ascertain if this provides a general method for determining binding sites of protein-bound oligosaccharides.

Acknowledgment. We thank Ms. Melissa J. Schur for preparation of enzymes, Dr. James L. Brunton for providing Verotoxin-1, Ms. Suzon Larocque for NMR technical help, and Mr. Don Krajcarski for MS analyses. We also thank Dr. James C. Richards for discussion and critical reading of the manuscript. This work is part of the Scientific Technical Cooperation Program between NRC (Canada) and CNRS (France). This is NRC paper 42446.

Supporting Information Available: ^{1}H and ^{13}C NMR spectra of **6**, **8**, **11**, and **14**, general procedure for enzymatic reaction, preparation of donor **8**, and glycosylation on MPEG. This material is available free of charge via the Internet at http://pubs.acs.org.

OL016466J

(14) (a) Fang, J.; Li, J.; Chen, X.; Zhang, X.; Wang, J.; Guo, Z.; Zhang, W.; Yu, L.; Brew, K.; Wang, P. G. *J. Am. Chem. Soc.* **1998**, *120*, 6635. (b) Hanessian, S.; Saavedra, O. M.; Mascitti, V.; Marterer, W.; Oehrlein, R.; Mak, C.-P. *Tetrahedron* **2001**, *57*, 3267.

(15) Schmidt, D.; Thiem, J. *J. Chem. Soc, Chem. Commun.* **2000**, 1919.

(16) Lubineau, A.; Malleron, A.; Le Narvor, C. *Tetrahedron Lett.* **2000**, *41*, 8887.

(17) FCHASE is a fluorescein derivative; see: Gilbert, M.; Cunningham, A.-M.; Watson, D. C.; Martin, A.; Richards, J. C.; Wakarchuk, W. W. *Eur. J. Biochem.* **1997**, *249*, 187.

(18) Clark, C.; Bast, D.; Sharp, A. M.; St. Hilaire, P.; Agha, R.; Stein, P. E.; Toone, E. J.; Read, R. J.; Brunton, J. L. *Mol. Microbiol.* **1996**, *19*, 891.

(19) (a) Shimizu, H.; Field, R. A.; Homans, S. W.; Donohue-Rolfe, A. *Biochemistry* **1998**, *37*, 11078. (b) Thompson, G. S.; Shimizu, H.; Homans, S. W.; Donohue-Rolfe, A. *Biochemistry* **2000**, *39*, 13153. (c) Ling, H.; Boodhoo, A.; Hazes, B.; Cummings, M. D.; Armstrong, G. D.; Brunton, J. L.; Read, R. J. *Biochemistry* **1998**, *37*, 1777.

(20) Yan, F. Y.; Wakarchuk, W. W.; Gilbert, M.; Richards, J. C.; Whitfield, D. M. *Carbohydr. Res.* **2000**, *328*, 3.