SYNTHESIS OF HUMAN CANCER ASSOCIATED GLOBO-H (MBR1) GLYCOSYLAMINO ACID: SOME MECHANISTIC AND CONFORMATIONAL REINVESTIGATIONS

Jianglong Zhu,[1] Qian Wan,[1] Guangbin Yang,[2] Ouathek Ouerfelli,[2] and Samuel J. Danishefsky*,[1,3]

[1] Laboratory for Bioorganic Chemistry, and [2] Organic Synthesis Core Laboratory, Sloan-Kettering Institute for Cancer Research, 1275 York Ave., New York, NY 10065, and [3] Department of Chemistry, Columbia University, 3000 Broadway, New York, NY 10027 (USA)

s-danishefsky@mskcc.org

Abstract – The synthesis of an extended globo-H (MBr1 antigen) in the form of a glycosylamino acid is reported. By careful NMR analysis, we found an interesting conformational "flip" on the E ring of some synthetic intermediates. An explanation offered for the successful [3+3] coupling of ABC acceptor **11** and DEF donor **10** possessing a C4 free hydroxyl to produce β-galactoside in azaglycosidations is reinforced.

In the 1980s, the glycosphingolipid antigen globo-H (**1**, Figure 1) was isolated by Hakomori and co-workers from the human breast cancer cell line MCF-7,[1] and was recognized by the monoclonal antibody MBr1.[2] This human breast cancer-associated antigen was later found to be overexpressed in a number of other carcinomas, including prostate, ovary, lung, colon, and small cell lung cancers.[3] Due to the complexity of this molecule, and our interest in developing carbohydrate-based anticancer vaccines,[4] we launched a program toward the total synthesis of the globo-H antigen. Indeed, the total synthesis of homogeneous globo-H was accomplished in our laboratory, and the synthetic route served to confirm the published structure of **1**.[5] This biologically interesting glycolipid has stimulated broad interest in the organic synthesis community, and subsequent syntheses have been accomplished by several groups, including Schmidt,[6] Boons,[7] Wong[8], Seeberger,[9] Huang,[10] and Wang.[11]

Our first-generation total synthesis of globo-H utilized all glycal building blocks for the rapid construction of the complex oligosaccharide.[5] These synthetic efforts relied on a critical and convergent [3+3] coupling, which made use of our sulfonamido glycosidation protocol[12] to generate the hexasaccharide core. In this route, a flexible terminal glycal was maintained throughout the hexasaccharide construction (Figure 1). Late stage appendage of the ceramide side chain or the allyl functional group to the glycal served to provide either globo-H glycolipid (**1**) or its allyl glycoside (**2**), respectively. The synthesis of **1** served to facilitate the proof of the stereostructure determination and the immunocharacterization of globo-H. The allyl glycoside **2** was employed for immunoconjugation to biocarrier proteins. This first generation synthetic route was sufficiently efficient to provide adequate quantities of synthetic material for proof of structure, immunocharacterization, conjugation, mouse vaccinations, and phase I clinical trials.
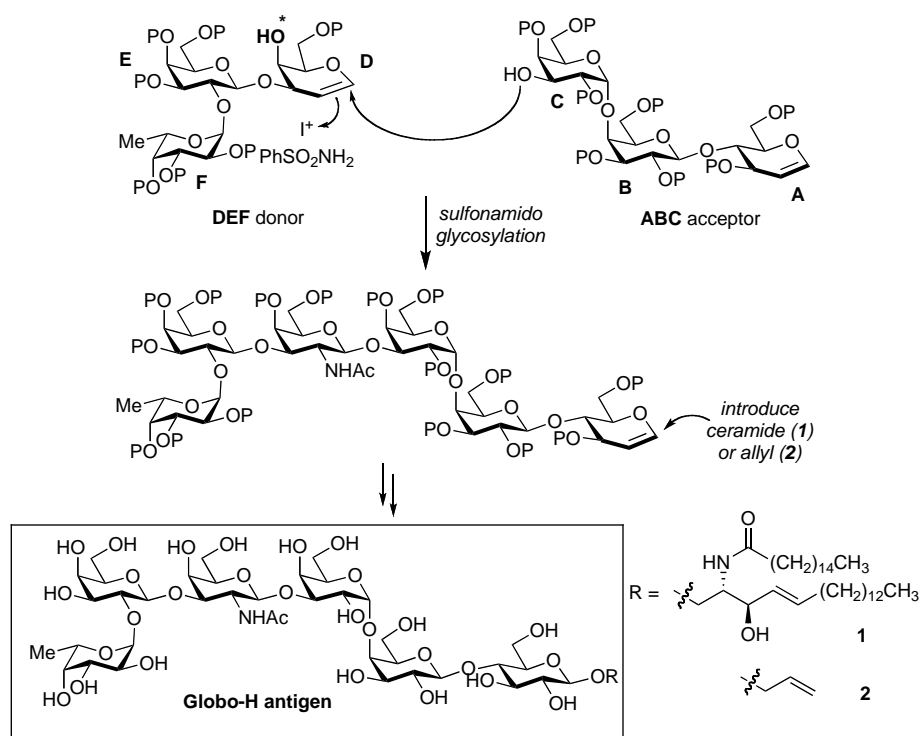

Figure 1. Our first generation synthesis of the human cancer associated globo-H antigen (MBr1).

When an anticancer vaccine based on this antigen was advanced to phase II/III clinical trials for prostate cancer in the late 1990s, it became necessary to develop a second generation synthesis of a modified globo-H antigen, incorporating an *n*-pentenyl, rather than allyl, glycoside.[13] In the hopes of achieving maximal convergency, we designed a modified ABC acceptor (**11**, Scheme 1), already equipped with the pentenyl glycosidic linker. The hexasaccharide core would then be assembled via an analogous [3+3] ABC+DEF coupling reaction. As previously described,[5] synthesis of DEF trisaccharide thiodonor (**10**) commenced with the preparation of the TIPS cyclic carbonate α-epoxide (**4**) through DMDO-mediated

epoxidation of cyclic carbonate (**3**).[14] This donor, in turn, was used to glycosylate **5** under the effect of zinc chloride, providing **6** in 87% yield. Next, upon exposure to fucosyl donor (**7**), under Mukaiyama[15]-Nicolaou[16] conditions, trisaccharide glycal (**8**) was obtained in 47% yield. The latter was then converted to the 1-*β*-ethanethiolate donor (**10**), bearing a 2-*α*-phenylsulfonamido function, via *trans*-diaxial iodosulfonamide (**9**) and thence to DEF donor **10,** drawing from our iodosulfonimidation/rearrangement methodology. Treatment of the DEF donor (**10**) with MeOTf in the presence of ABC acceptor (**11**) smoothly provided hexasaccharide (**12**) in 60-70% yield. Following deprotection, Birch reduction, and peracetylation, **12** was advanced to the peracetylated hexasaccharide (**13**), in 42% overall yield. After de-aceylation, the terminal pentenyl function of the globo-H glycoside was subjected to oxidative cleavage, and the resulting *γ*-oxybutyryl aldehyde was conjugated to KLH *via* reductive amination.[13] In addition, **13** was later employed to prepare the glycosylamino acid (**14**, Scheme 1).[17]
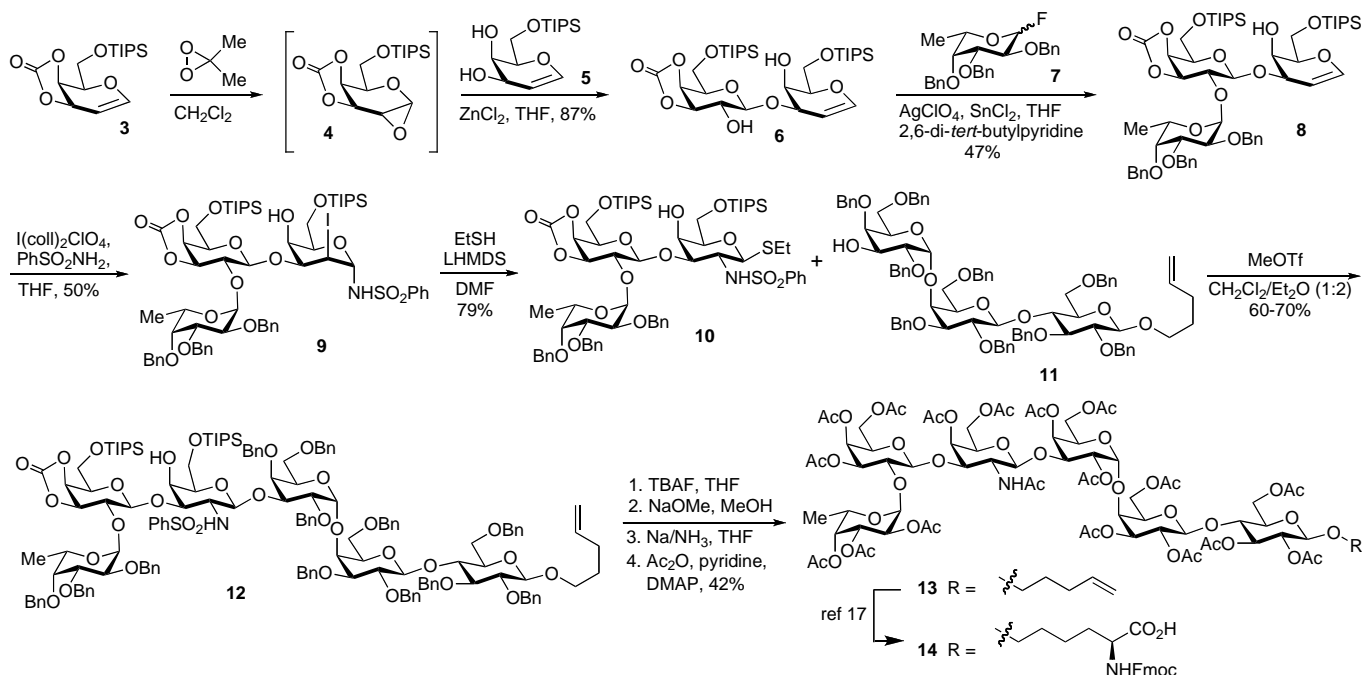

Scheme 1. Our previous synthesis of globo-H pentenyl glycoside and glycosylamino acid.

In the context of our ongoing, carbohydrate-based vaccine program, we recently synthesized a unimolecular pentavalent construct, incorporating five different carbohydrate-based antigens – including globo-H – whose overexpression is associated with breast and prostate cancer.[18] We are currently advancing this unimolecular pentavalent construct to phase I clinical trials. Toward that end, we recently scaled up the synthesis of the globo-H antigen, according to the route outlined in Scheme 1.[19]

NMR Studies. During the course of the synthesis of the globo-H antigen, we observed some interesting conformational changes on the D-galactose moiety of **6**, **8**, **10**, **12**, **13** and **15** (E-ring, colored in red, Scheme 2).   These phenomena were observed through careful examination of the coupling constants between vicinal hydrogen atoms ($J_{H1,H2}$) in extensive 2D NMR studies. Thus, we observed that the cyclic carbonate and TIPS-protected D-galactose in disaccharide (**6**) adopts a regular stable $^4C_1$ conformation (E-ring, $J_{H1,H2}$ = 6.6 Hz). However, when the L-fucose moiety was appended to the disaccharide (**6**), an interesting ring flip took place on the D-galactose moiety of the resultant trisaccharide (**8**) (Scheme 2). The small coupling constant observed between the vicinal hydrogen atoms of the D-galactose moiety of **8** (E ring, $J_{H1,H2}$ = 4.8 Hz) is in agreement with a distorted $^1C_4$ conformation (see **8'**).[20] This ring flip may be attributed to the steric congestion of the E-ring, which presents a β-D-Galactal, an α-D-Fucose (1→2), a 3, 4-cyclic carbonate and a 6-*O*-TIPS ether. Unfavorable steric interactions of the D-galactose moiety would presumably be somewhat alleviated in this distorted $^1C_4$ conformation.

The 1-β-ethanethiolate donor (**10**) prepared from **8** also adopts a $^1C_4$ conformation at the D-galactose moiety (see **10'**, E-ring, $J_{H1,H2}$ = 1.0 Hz). Moreover, the smaller $J_{H1,H2}$ in the E-ring of **10** indicate that it exists in a more standard $^1C_4$ conformation, which is consistent with the relatively larger size of the D-ring of **10** (a protected β-D-galactose amine) in comparison with that of **8** (a β-D-galactal). Along these lines, one would expect that the hexasaccharide (**12**), which possesses the entire protected ABCDEF system, should adopt an even more regular $^1C_4$ conformation at the E-ring (colored in red). Indeed, this conjecture is in complete agreement with the observed coupling constant between the vicinal hydrogen atoms at the E-ring of **12** ($J_{H1,H2} \approx 0$ Hz). With the removal of the two –TIPS protecting groups and the cyclic carbonate, the resulting hexasaccharide (**15**)[21] now assumed a regular stable $^4C_1$ conformation at the E-ring. Hexasaccharide (**15**) was further subjected to dissolving metal reduction, resulting in the removal of all benzyl and sulfonamide protecting groups, followed by peracetylation to furnish the isolable hexasaccharide peracetate (**13**).   The latter was also observed to adopt a regular stable $^4C_1$ conformation at the E-ring (Scheme 2). The change in conformation from $^1C_4$ to $^4C_1$ may be attributed to the release of the steric interactions on the D-galactose moiety. HSQC analysis on **15** and **13** clearly shows that these compounds possess two α- and four β-anomeric centers, which is consistent with the natural globo-H ceramide (MBr1 antigen).
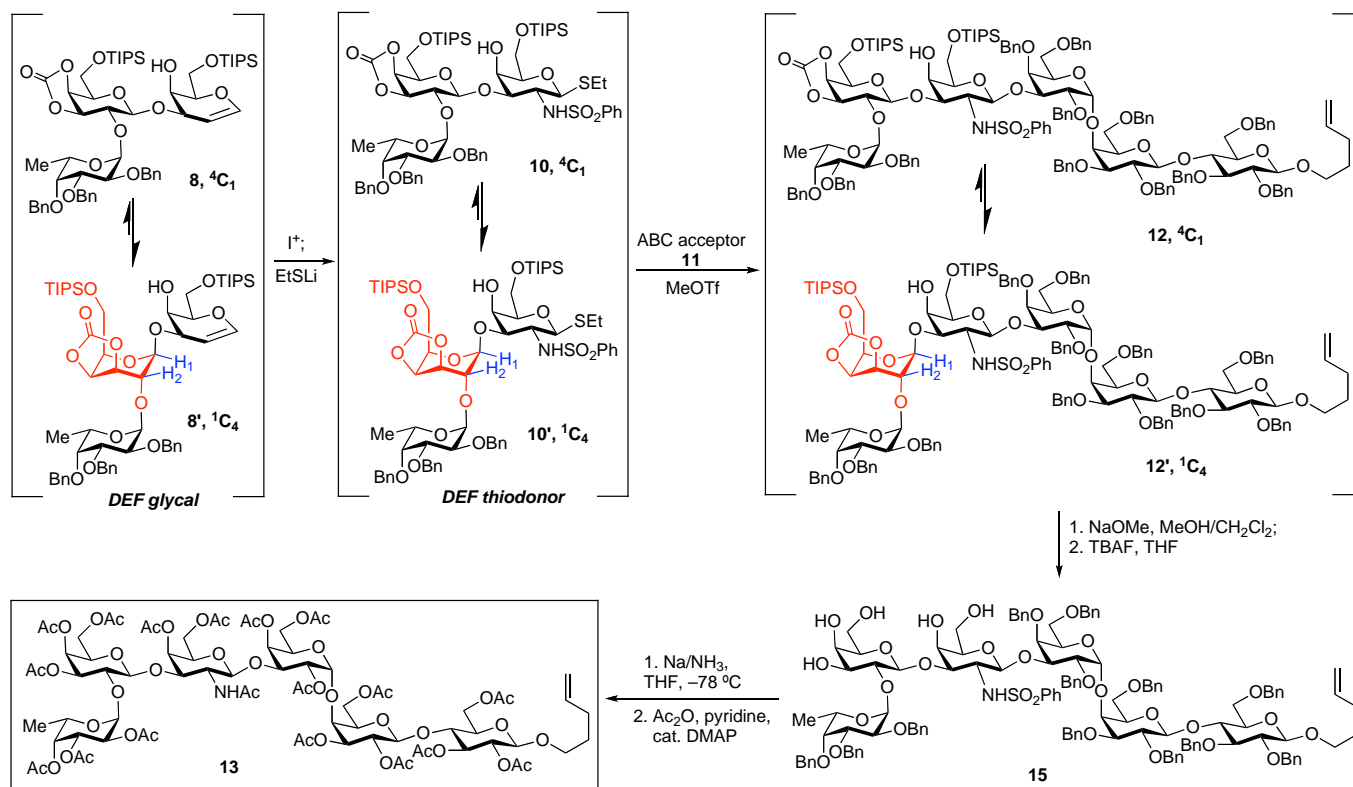
Scheme 2. Conformational reinvestigation.

Mechanistic Analysis of the [3+3] Coupling Reaction. Key to the success of the [3+3] coupling of ABC acceptor (**11**) and DEF donor (**10/10'**) is the use of galactosyl donor carrying a free hydroxyl group at C4. As we previously reported, the preferential formation of *β*-galactoside in aza-glycosidations with galactosyl donors possessing a C4 free hydroxyl group appears to be quite general.[5,22] By contrast, when C4 is protected as an acetate or benzyl ether, such as in **19** and **20**, competition from *α*-azagalactoside formation can be severe.[5,23] We had previously reported a detailed explanation for this interesting stereochemical outcome based on the observed experimental data.[23] We present herein some further insights into this observed phenomenon. According to the coupling constants analysis described above, the D-galactose thiodonor moiety of the DEF donor adopts a stable $^4C_1$ conformation. Upon activation of ethyl thioether with methyl triflate, conventional neighboring group participation of the C_2 sulfonamide could possibly lead to one of two structures: the five-membered intermediate (**16**), derived from oxygen attack, or the three-membered compound (**17**), resulting from nitrogen attack (Scheme 3). Due to possible intramolecular hydrogen bonding between the C4-hydroxyl and the pyranosidal oxygen (see reference 19 for corresponding data analysis), participation from the ring oxygen to expel an equatorial anomeric group, *en route* to an oxonium intermediate **18**, would appear to be electronically disfavored. Thus, nucleophilic attack of the ABC acceptor on either **16** or **17** would give rise to the *β*-galactoside product, (i.e. **12/12'**). Along these same lines, activation of the ethyl thioether of either **19** or **20** with methyl

triflate would lead to either **21**, derived from oxygen attack, or **22**, resulting from nitrogen attack. However, when C4 is protected as an acetate (**19**) or benzyl ether (**20**), the absence of intramolecular hydrogen bonding would better dispose the ring oxygen to expel the equatorial anomeric group in **21** or **22**, *en route* to an oxonium intermediate (**23**).   Subsequent nucleophilic attack of the ABC acceptor to **23** would lead to the formation of α-galactoside as the anomerically favored product. These experimental results indicate that subtle changes in the protecting groups could dramatically affect the glycosyl donor conformations, and subsequently lead to different stereochemical outcomes of glycosidation.[23]
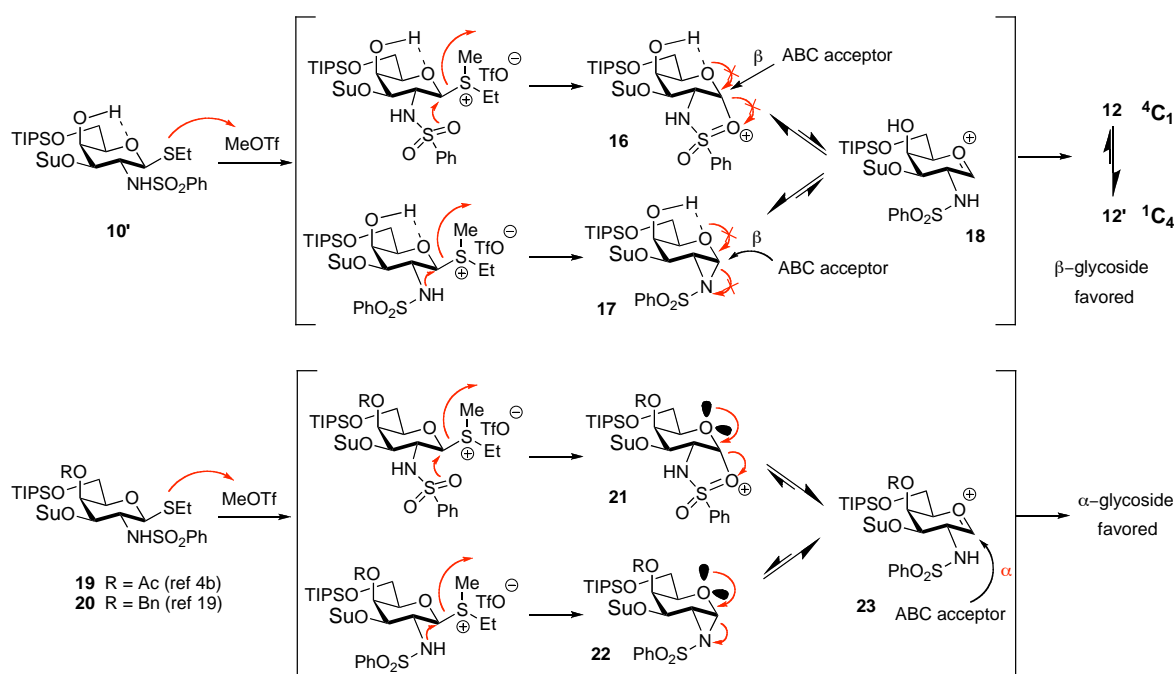
Scheme 3. Mechanistic proposal for the azaglycosidations.

Synthesis of Modified Glycosyl Amino Acid.   In order to facilitate the preclinical and clinical evaluation of the unimolecular pentavalent vaccine construct targeting prostate and breast cancer bearing Globo-H, GM2, STn, TF and Tn,[18] we carried out the synthesis of chain-extended Globo-H glycosylamino acid (**26**). The logic relied was based on our previously reported cross metathesis method,[24] with slight modifications. Thus, as shown in Scheme 4, cross-metathesis of hexasaccharide peracetate **13** with allylglycine benzylester (**24**), using the Grubbs-Hoveyda first generation catalyst (**25**)[25] and subsequent hydrogenative reduction afforded the desired chain-extended Globo-H glycosylamino acid (**26**)[26] in 58% yield over two steps.
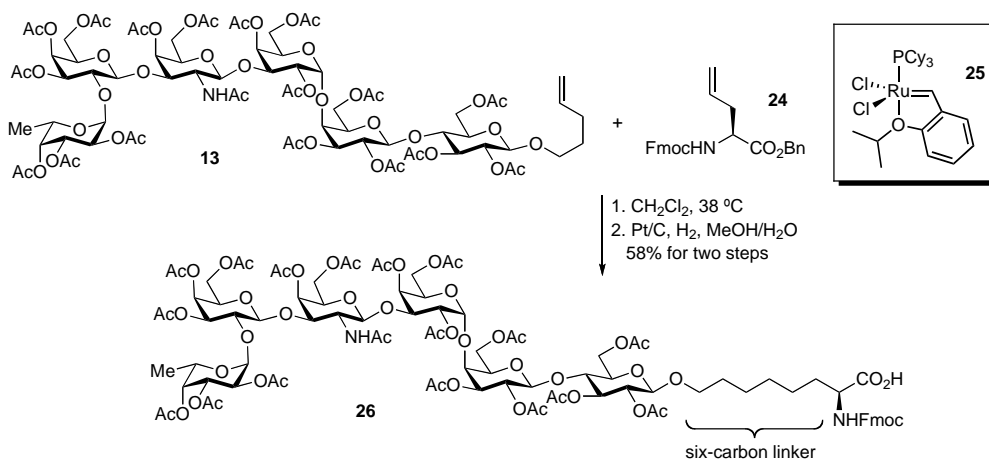
Scheme 4. Synthesis of a chain-extended globo-H glycosylamino acid.

REFERENCES AND NOTES

1. R. Kannagi, S. B. Levery, F. Ishijamik, S. Hakomori, L. H. Schevinsky, B. B. Knowles, and D. Solter, *J. Biol. Chem.*, 1983, **258**, 8934; E. G. Bremer, S. B. Levery, S. Sonnino, R. Ghidoni, S. Canevari, R. Kannagi, and S. Hakomori, *J. Biol. Chem.*, 1984, **259**, 14773.

2. S. Menard, E. Tagliabue, S. Canevari, G. Fossati, and M. I. Colnaghi, *Cancer Res.*, 1983, **43**, 1295.

3. P. O. Livingston, *Semin. Cancer Biol.*, 1995, **6**, 357.

4. R. M. Wilson and S. J. Danishefsky, "Fully Synthetic Carbohydrate-Based Antitumor Vaccines," in *Cancer Vaccines*, 2007, ed. by R. Orentas, *In Press*.

5. M. T. Bilodeau, T. K. Park, S. Hu, J. T. Randolph, S. J. Danishefsky, P. O. Livingston, and S. Zhang, *J. Am. Chem. Soc.*, 1995, **117**, 7840; T. K. Park, I. J. Kim, S. Hu, M. T. Bilodeau, J. T. Randolph, O. Kwon, and S. J. Danishefsky, *J. Am. Chem. Soc.*, 1996, **118**, 11488.

6. J. Lassaletta and R. R. Schmidt, *Liebigs Ann.*, 1996, 1417.

7. T. Zhu and G.-J. Boons, *Angew. Chem. Int. Ed.*, 1999, **38**, 3495.

8. F. Burkhart, Z. Zhang, S. Wacowich-Sgarbi, and C.-H. Wong, *Angew. Chem. Int. Ed.*, 2001, **40**, 1274.

9. F. Bosse, L. A. Marcaurelle, and P. H. Seeberger, *J. Org. Chem.*, 2002, **67**, 6659; D. B. Werz, B. Castagner, and P. H. Seeberger, *J. Am. Chem. Soc.*, 2007, **129**, 2770.

10. Z. Wang, L. Zhou, K. El-Boubbou, X.-S. Ye, and X. Huang, *J. Org. Chem.*, 2007, **72**, 6409.

11. D. M. Su, H. Eguchi, W. Yi, L. Li, P. G. Wang, and C. Xia, *Org. Lett.*, 2008, **10**, 1009.

12. D. A. Griffith and S. J. Danishefsky, *J. Am. Chem. Soc.*, 1990, **112**, 5811; D. A. Griffith and S. J. Danishefsky, *J. Am. Chem. Soc.*, 1991, **113**, 5863.

13. J. R. Allen, J. G. Allen, X.-F. Zhang, L. J. Williams, A. Zatorski, G. Ragupathi, P. O. Livingston, and S. J. Danishefsky, *Chem. Eur. J.*, 2000, **6**, 1366.

14. R. L. Halcomb and S. J. Danishefsky, *J. Am. Chem. Soc.,* 1989, **111**, 6661.

15. T. Mukaiyama, Y. Murai, and S. Shoda, *Chem. Lett.*, 1981, 431.

16. K. C. Nicolaou, T. J. Caulifield, H. Kataoka, and N. A. Stylianides, *J. Am. Chem. Soc.*, 1990, **112**, 3693; K. C. Nicolaou, C. W. Hummel, N. J. Bockovich, and C.-H. Wong, *J. Chem. Soc., Chem. Commun.*, 1991, 870.

17. J. R. Allen, C. R. Harris, and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2001, **123**, 1890.

18. G. Ragupathi, F. Koide, P.O. Livingston, Y. S. Cho, E. Atsushi, Q. Wan, M. K. Spassova, S. J. Keding, J. Allen, O. Ouerfelli, R. M. Wilson, and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2006, **128**, 2715.

19. Detailed procedures for preparation of all compounds except **15** and **26** and their characterization data have previously been reported in the literature (refs 5 and 13).

20. Y. Rao and G.-J. Boons, *Angew. Chem. Int. Ed.*, 2007, **46**, 6148.

21. Synthesis of **15**: To hexasaccharide (**12**, 215 mg, 0.083 mmol) in a mixture of 4 mL MeOH and 6 mL CH_2Cl_2 was added NaOMe (0.5 M in MeOH, 1.3 mL), and the reaction mixture was stirred at rt overnight. The solvents were removed under reduced pressure, the residue was then diluted with EtOAc, and sequentially washed with saturated aqueous ammonium chloride and brine. The organic layer was separated, dried with sodium sulfate, filtered, and concentrated. The residue was dissolved in 5 mL THF, TBAF (1.0 M in THF, 0.66 mL) was added, the reaction was stirred at rt overnight. The reaction mixture was concentrated and purified on silica gel column chromatography (CH_2Cl_2 to 5% MeOH in CH_2Cl_2) to afford 183 mg **15** (98% yield for two steps). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.70 (2H, d, J = 7.4 Hz), 7.39~6.89 (63H, m), 5.75~5.66 (2H, ovrlp), 5.26 (1H, d, J = 11.6 Hz), 5.15 (1H, br), 4.93~4.85 (5H, ovrlp), 4.78~4.59 (10H, ovrlp), 4.51~4.33 (6H, ovrlp), 4.26~4.17 (5H, ovrlp), 4.09~4.01 (6H, ovrlp), 3.97 (1H, br), 3.91~3.75 (12H, ovrlp), 3.70 (2H, m), 3.63~3.40 (13H, ovrlp), 3.27~3.06 (7H, ovrlp), 2.56 (1H, m), 2.05 (2H, m), 1.64 (2H, m), 0.73 (3H, d, J = 6.4 Hz); ^{13}C NMR (150 MHz, $CDCl_3$) δ 142.35, 139.56, 139.47, 138.98, 138.74, 138.68, 138.55, 138.53, 138.35, 138.31, 138.26, 138.05, 137.95, 137.15, 131.62, 128.97, 128.72, 128.51, 128.43, 128.38, 128.33, 128.30, 128.27, 128.22, 128.17, 128.14, 128.12, 128.09, 128.01, 127.96, 127.89, 127.75, 127.63, 127.55, 127.52, 127.49, 127.43, 127.40, 127.38, 127.34, 127.30, 127.27, 127.17, 127.07, 127.01,

126.98, 103.47, 102.84, 102.68, 100.55, 100.36, 100.08, 82.64, 81.47, 81.43, 81.39, 79.19, 78.95, 78.23, 78.09, 77.48, 77.34,   77.15, 75.24, 75.14, 75.11, 75.00, 74.97, 74.79, 74.70, 74.67, 74.65, 73.27, 73.16, 73.12, 72.99, 72.65, 72.57, 71.95, 69.60, 69.19, 68.52, 68.32, 68.06, 67.99, 67.53, 67.40, 62.38, 61.84, 54.80, 30.17, 28.90, 16.73; ESI-MS m/z 2285.1 [M+Na]$^+$; 1154.1 [M+2Na]$^{2+}$.

22. S. J. Danishefsky and M. T. Bilodeau, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 1380.

23. O. Kwon and S. J. Danishefsky, *J. Am. Chem. Soc.*, 1998, **120**, 1588.

24. Q. Wan, Y. S. Cho, T. H. Lambert, and S. J. Danishefsky, *J. Carbohydr. Chem.*, 2005, **24**, 425.

25. J. S. Kingsbury, J. P. A. Harrity, P. J. Bonitatebus, Jr., and A. H. Hoveyda, *J. Am. Chem. Soc.*, 1999, **121**, 791.

26. Synthesis of **26**: To an oven-dried round-bottom flask containing peracetate (**13**, 88 mg, 0.049 mmol) and allylglycine benzyl ester (**24**, 168 mg, 0.39 mmol) in 1.0 mL anhydrous CH_2Cl_2 was added Grubbs-Hoveyda first generation catalyst (**25**, 7.2 mg, 0.12 mmol) under argon. The flask was then equipped with a finger condenser, and the reaction mixture was stirred at 38 ℃ until full disappearance of **13**. Concentration followed by purification on silica gel column chromatography (1/1 hexanes/EtOAc→100% EtOAc→5% MeOH in EtOAc) gave the crude desired product. This crude material was dissolved in a mixture of 4.5 mL MeOH and 0.3 mL water, 19 mg of 10% Pt/C was added. The reaction mixture was stirred under an atmosphere of hydrogen for three to four days. The mixture was concentrated, purified first on silica gel column chromatography (100% EtOAc→5% MeOH in EtOAc), then by preparative TLC (25/1 EtOAc/ MeOH, with 2% HOAc) to afford 60 mg desired product (**26**, 58% for two steps). ^{1}H NMR (600 MHz, CD$_3$OD) δ 7.81 (2H, d, J = 7.5 Hz), 7.69 (2H, t, J = 8.8 Hz), 7.40 (2H, t, J = 7.5 Hz), 7.32 (2H, t, J = 7.3 Hz), 5.65 (1H, d, J = 2.4 Hz), 5.44 (1H, d, J = 2.8 Hz), 5.31 (1H, d, J = 2.8 Hz), 5.26 (1H, d, J = 3.1 Hz), 5.22~5.15 (4H, ovrlp), 5.10~5.06 (2H, m), 4.97~4.80 (5H, ovrlp), 4.70 (2H, t, J = 5.9 Hz), 4.63 (1H, d, J = 7.9 Hz), 4.57 (1H, d, J = 8.0 Hz), 4.49~4.36 (6H, ovrlp), 4.25~3.94 (17H, ovrlp), 3.85~3.80 (3H, ovrlp), 3.72~3.66 (2H, ovrlp), 3.50 (1H, m), 2.18~1.94 (58H, ovrlp), 1.84~1.28 (10H, ovrlp), 1.16 (3H, d, J = 6.4 Hz); ^{13}C NMR (150 MHz, CD$_3$OD) δ 172.31, 172.12, 172.11, 172.08, 172.02, 171.91, 171.88, 171.80, 171.67, 171.64, 171.42, 171.26, 158.80, 145.55, 145.38, 142.75, 129.00, 128.38, 128.35, 126.47, 121.13, 103.88, 102.97, 102.27, 101.79, 100.76, 97.61, 78.35, 78.19, 75.79, 75.01, 74.68, 74.63, 74.21, 74.01, 73.58, 73.40, 72.94, 72.62, 72.43, 71.99, 71.70, 71.51, 71.00, 69.99, 69.10, 69.01, 68.92, 68.05, 66.46, 63.85, 63.79, 63.37, 62.86, 62.71, 48.61, 32.82, 30.57, 29.87, 27.02, 26.96, 23.69, 21.84, 21.47, 21.27, 21.24, 21.22, 21.14, 21.10, 20.98, 20.95, 20.91, 20.85, 20.77, 20.73, 20.67, 20.63, 16.41; ESI-MS m/z 2131.9 [M+Na]$^+$; 1077.5 [M+2Na]$^{2+}$.