# Synthesis of a Tumor-Associated 2,3-Sialyl-T Glycododecapeptide Antigen from the Tandem Repeat Region of the Mucin MUC1

Sebastian Dziadek, Horst Kunz*

Institut für Organische Chemie, Johannes Gutenberg-Universität Mainz, Duesbergweg 10-14, 55128 Mainz, Germany
Fax +49(6131)3924786; E-mail: hokunz@uni-mainz.de

*Received 18 June 2003*

Dedicated to Professor Wolfgang Steglich on the occasion of his 70th birthday.

**Abstract:** As a cell surface antigen for the development of selective antitumor vaccines, a tumor-associated glycododecapeptide from epithelial MUC1 carrying the 2,3-sialyl-T antigen was synthesized. The 2,3-ST building block was assembled from a $T_N$-threonine conjugate by stepwise saccharide chain extension and utilized for the solid-phase glycopeptide synthesis on the fluoride-labile PTMSEL linker.

**Key words:** 2,3-sialyl-T antigen, glycopeptides, MUC1, regioselective sialylation, solid-phase synthesis

In contrast to normal cells, tumor cells are significantly altered in their glycoprotein profile on the outer cell surface.[1] This difference constitutes the basis for a selective immunological attack at certain cancer cells.

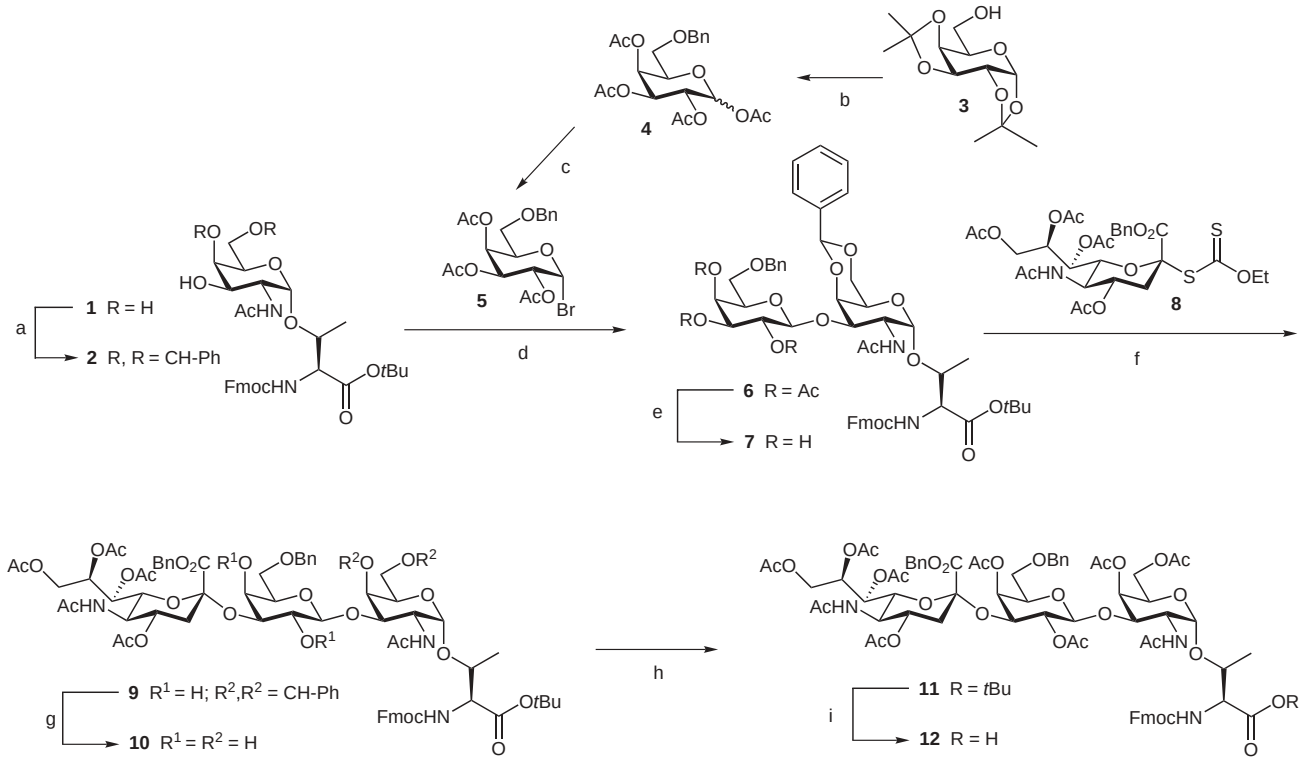
The epithelial mucin MUC1[2] is a heavily *O*-glycosylated glycoprotein present at the interface between many epithelia and their extracellular environments.[3] The extracellular domain consists of tandem repeats comprising 20 amino acids of the sequence HGVTSAPDTRPAPGSTAPPA, including five potential O-glycosylation sites. In epithelial tumor cells, the expression of MUC1 is drastically increased which is combined with an incomplete formation of the glycan side-chains due to a premature sialylation. The aberrant glycosylation results in the exposure of additional peptide epitopes which hence become accessible to the immune system. The 2,3-sialyl-T antigen (αNeuNAc-2,3-βGal-1,3-αGalNAc-*O*-Ser/Thr) constitutes an important tumor-associated saccharide which was found on breast cancer cells.[4] Glycopeptides, containing partial structures of cancer-associated cell-surface glycoproteins, are considered promising candidates for the construction of tumor-selective immunostimulating antigens. In particular, glycopeptides from MUC1 have been shown to be processed by antigen-presenting dendritic cells and presented via MHC molecules – leaving the glycan side-chain intact[5] – thus bearing the potential of inducing both humoral as well as cellular immune responses. For the development of antitumor vaccines, the generation of synthetic glycopeptides, combining both saccharide as well as peptide epitopes, is therefore a feasible strategy.

We here describe a novel synthetic approach toward the tumor-associated 2,3-sialyl-T antigen building block and its incorporation into a glycododecapeptide of the tandem repeat domain of the MUC1 mucin. In contrast to other strategies,[6] the complex sialic acid was introduced at a late stage of the synthesis in a regio- and stereoselective reaction. Starting from the readily available glycosylated amino acid Fmoc-Thr(αGalNAc)-O-*t*-Bu[7] **1**, the T disaccharide was synthesized first and subsequently converted into the α-2,3-sialyl T antigen (Scheme 1).

For the 3-β-galactosylation of Fmoc-Thr(αGalNAc)-O-*t*-Bu **1,** protection of the 4- and 6-OH groups as benzylidene acetal was necessary to give **2** (yield: 83%).[8] The subsequent glycosylation using 6-*O*-benzyl protected galactosyl bromide **5** activated with $Hg(CN)_2$[9] in a mixture of nitromethane/dichloromethane (3:2) furnished the disaccharide conjugate **6** in a yield of 85%. Galactosyl bromide **5** was prepared from 1,2:3,4-diisopropylidene galactose **3**, which was subsequently benzylated at the 6-OH function (87%). Following acidolysis of the acid-labile isopropylidene protecting groups, acetylation (**4**, 89% over two steps) and careful treatment with HBr in glacial acid afforded 55% of the galactosyl donor **5** after silica gel filtration.[10]

Due to steric hindrance, O-deacetylation of the conjugate **6** proved difficult. Only careful, prolonged treatment with NaOMe/MeOH at pH = 8.5 furnished the acceptor **7** in a yield of 62%. The key step of the synthetic strategy involved the stereo- and regioselective sialylation of the T antigen disaccharide **7** at the 3′-hydroxy group. For this purpose, xanthate **8** of the *N*-acetylneuraminic acid benzyl ester[11] was activated with phenylsulfenyl triflate, generated in situ from phenylsulfenyl chloride and silver triflate.[12] Di-*tert*-butylpyridine (DTBP) served as proton scavenger.[12] Low temperature and the use of acetonitrile/dichloromethane (2:1) as solvent favored the stereoselective formation of the α-2,3-sialyl-T threonine conjugate **9** in a yield of 50% (α/β, 97:3).[13] Cleavage of the benzylidene acetal with acetic acid (**10**, 82%) and O-acetylation with pyridine/acetic anhydride (2:1) catalyzed by *N*,*N*-dimethylamino pyridine (DMAP)[14] afforded the protected compound **11** in a yield of 87%. Subsequent careful acidolysis of the *tert*-butyl ester quantitatively furnished the Fmoc-2,3-ST threonine building block **12**.[15]

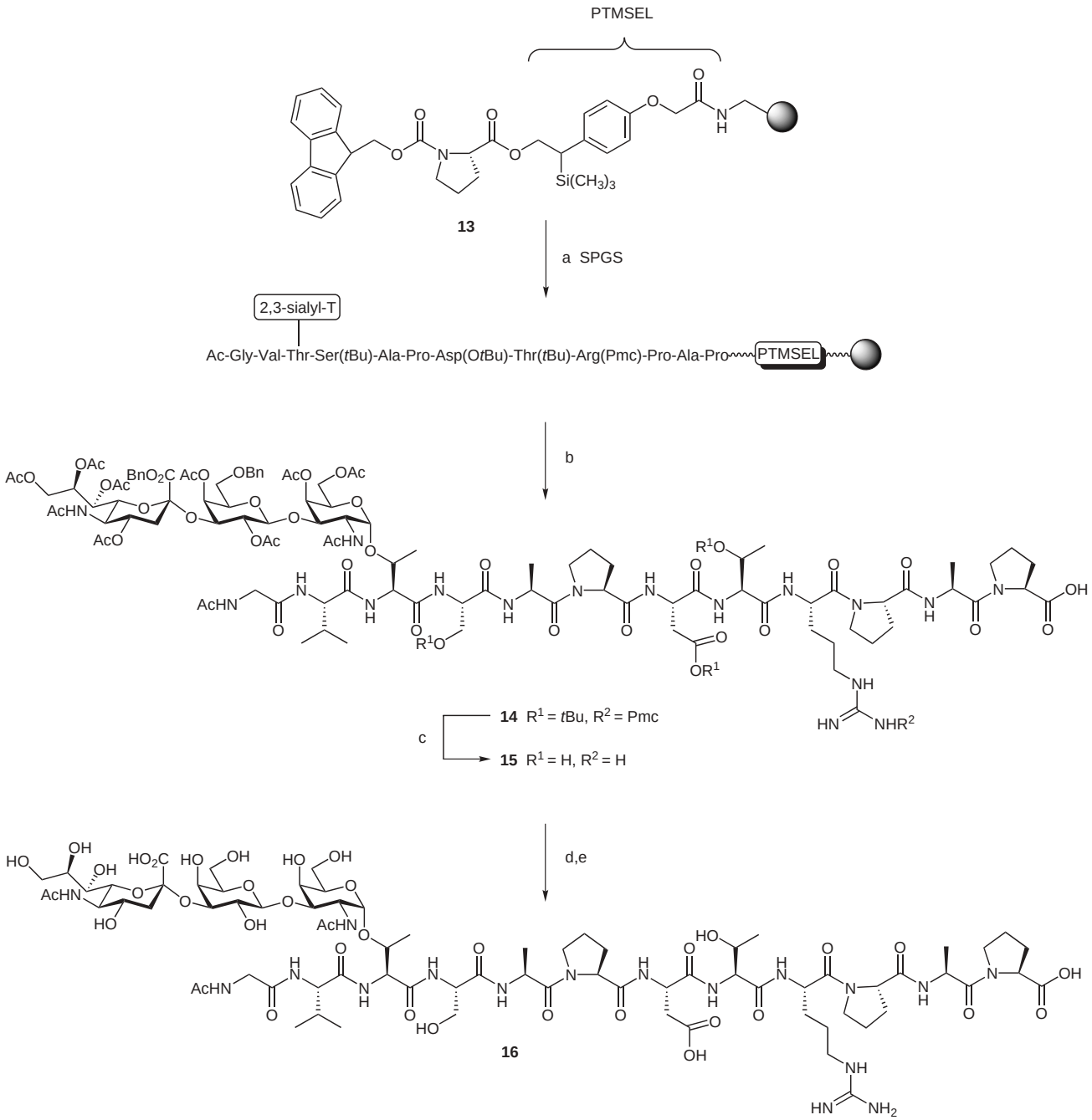
Synlett 2003, No. 11, Print: 02 09 2003. Web: 11 08 2003.
Art Id.1437-2096,E;2003,0,11,1623,1626,ftx,en;G13903ST.pdf.
DOI: 10.1055/s-2003-41009



**Scheme 1** *Reagents and conditions*: a) α,α-dimethoxytoluene, *p*-TsOH, $CH_3CN$, 3 h, 83%. b) i) NaH, BnBr, DMF, 24 h, 87%; ii) 1. AcOH (60%), 85 °C, 2. pyridine/$Ac_2O$ (2:1), DMAP, 6 h, 89%. c) HBr in AcOH (33%), $CH_2Cl_2$, 0 °C, 30 min, 55%. d) $Hg(CN)_2$, $CH_3NO_2/CH_2Cl_2$ (3:2), MS 3 Å, 25 h, 85%. e) NaOMe/MeOH, pH = 8.5, 12 h, 62%. f) $CH_3CN/CH_2Cl_2$ (2:1), MS 3Å, PhSCl, AgOTf, DTBP, –68 °C, 4 h, 50% (α/β, 97:3). g) AcOH (80%), 80 °C, 1 h, 82%. h) pyridine/$Ac_2O$ (2.3:1), DMAP, 6h, 87%. i) TFA, anisole, quant. DTBP = di-*tert*-butylpyridine.

The solid-phase synthesis of a glycododecapeptide from the tandem repeat region of MUC1 carrying the 2,3-sialyl-T antigen was performed in a peptide synthesizer[16] according to the Fmoc-strategy. NovaSyn Tenta Gel resin equipped with the novel fluoride-labile PTMSEL linker[17] was employed and loaded with the starting amino acid Fmoc-proline. This anchor can be cleaved under almost neutral conditions by applying tetrabutylammonium fluoride trihydrate (TBAF·3 $H_2O$) in dichloromethane. Under these conditions aspartate structures do not rearrange via aspartimides, and most of the common protecting groups in Fmoc-peptide chemistry are not affected. The first eight amino acids of the glycododecapeptide were coupled according to the standard protocol. After removal of the Fmoc-protecting group of **13** with piperidine in *N*-methylpyrrolidone (NMP), the extension of the peptide chain was performed by iterative coupling using a 10-fold excess of each amino acid activated by HBTU[18]/HOBt and diisopropylethylamine (DIPEA) in NMP (Scheme 2). Unreacted amino groups were capped after each step with $Ac_2O$/DIPEA/HOBt (4:1:0.12) in NMP. The 2,3-sialyl-T threonine building block **12** (1.4 equiv) was coupled manually over a period of 4 h using HATU/HOAt[19] and *N*-methylmorpholine (NMM) for activation. The final two amino acids were attached according to the standard protocol. After completed synthesis the *N*-terminal Fmoc group was exchanged for an acetyl group. The glycoconjugate was detached from the resin by treatment with 2 × 2.5 equivalents TBAF·3$H_2O$ in dichloromethane (45 min each) to furnish the protected glycopeptide **14** in a yield of 39% (based on the loaded resin **13**) after purification by semi-preparative RP-HPLC. Subsequently, the acid-labile side-chain protecting groups were removed by acidolysis with a mixture of trifluoroacetic acid, triisopropylsilane and water (15:0.9:0.9). The partially deprotected glycoconjugate **15** was precipitated into diethyl ether and thoroughly washed. Cleavage of the sialic acid benzyl ester as well as the galactose benzyl ether was accomplished by hydrogenolysis (23 h) in methanol using palladium on activated charcoal (10%) as catalyst. Complete final deacetylation was only achieved by treatment with aqueous NaOH (5 mM). At first, a derivative occurred as the sole product still carrying one acetyl group, presumably at the hindered 4′-OH group of the galactose portion. Extended reaction time slowly furnished the desired product. The reaction was carefully monitored by analytical HPLC. After 59 h the reaction was terminated with a conversion of 89%. Neither β-elimination of the glycan portion nor epimerization were detectable. The pure, completely deprotected glycododecapeptide **16**[20] was obtained after purification by semipreparative RP-HPLC in a yield of 56% over three steps.

The immunological evaluation of this synthetic 2,3-sialyl-T glycopeptide is currently in progress.




**Scheme 2** *Reagents and conditions*: a) solid-phase glycopeptide synthesis (SPGS): *Fmoc-removal* [piperidine/NMP (20%)], *coupling* (1.-8., 10.-11.: Fmoc-AA-OH (10 equiv), HBTU/HOBt/DIPEA, DMF, 9. **12** (1.4 equiv), HATU/HOAt/NMM, DMF, 4 h), *capping*: $Ac_2O$, DIPEA, HOBt (4:1:0.12). b) TBAF·$3H_2O$ (2 × 2.5 equiv), $CH_2Cl_2$, 45 min (each time), 39% (with respect to **13**). c) TFA/TIS/$H_2O$ (15:0.9:0.9), 2 h. d) $H_2$, Pd/C (10%), MeOH, 23 h. e) $NaOH_{aq}$ (5 mM), 59 h (56% over 3 steps). HATU = *O*-(7-aza-benzotriazole-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate, HBTU = *O*-(1*H*-benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate, HOAt = *N*-hydroxy-7-azabenzotriazole, HOBt = *N*-hydroxybenzotriazole; NMM = *N*-methylmorpholine; TBAF = tetrabutylammonium fluoride; TIS = triisopropylsilane; Pmc = 2,2,5,7,8-pentamethylchroman-6-sulfonyl-; PTMSEL = (2-phenyl-2-trimethylsilyl)ethyl-linker.

## Acknowledgment

This work was supported by the Deutsche Forschungsgemeinschaft and by the Stiftung Rheinland-Pfalz für Innovation. S. D. is grateful for a doctoral fellowship provided by the Fonds der Chemischen Industrie and the Bundesministerium für Bildung und Forschung (BMBF).

## References

(1) (a) Feizi, T. *Nature (London)* **1985**, *314*, 53. (b) Springer, G. F. *Science* **1984**, *224*, 1198.
(2) (a) Taylor-Papadimitriou, J.; Burchell, J. M.; Plunkett, T.; Graham, R.; Correa, I.; Miles, D.; Smith, M. *J. Mammary Gland Biol. Neoplasia* **2002**, *2*, 209. (b) Hanisch, F.-G.;

Müller, S. *Glycobiology* **2000**, *10*, 439. (c) von Mensdorff-Pouilly, S.; Snijdewint, F. G. M.; Verstraeten, A. A.; Verheijen, R. H. M.; Kenemans, P. *Int. J. Biol. Markers* **2000**, *15*, 343. (d) Taylor-Papadimitriou, J.; Burchell, J. M.; Miles, D. W.; Dalziel, M. *Biochim. Biophys. Acta* **1999**, 301.

(3) (a) Gum, J. R. *Biochem. Soc. Trans.* **1995**, *23*, 795. (b) Carlstedt, I.; Davies, J. R. *Biochem. Soc. Trans.* **1997**, *25*, 214.

(4) Brockhausen, I.; Yang, J.-M.; Burchell, J.; Whitehouse, C.; Taylor-Papadimitriou, J. *Eur. J. Biochem.* **1995**, *233*, 607.

(5) Vlad, A. M.; Müller, S.; Cudic, M.; Paulsen, H.; Otvos, L. Jr.; Hanisch, F.-G.; Finn, O. J. *J. Exp. Med.* **2002**, *11*, 1435.

(6) (a) Nakahara, Y.; Nakahara, Y.; Ito, Y.; Ogawa, T. *Carbohydr. Res.* **1998**, *309*, 287. (b) Komba, S.; Meldal, M.; Werdelin, O.; Jensen, T.; Bock, K. *J. Chem. Soc., Perkin Trans. 1* **1999**, 415. (c) Schwarz, J. B.; Kuduk, S. D.; Chen, X.-T.; Sames, D.; Glunz, P. W.; Danishefsky, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 2662.

(7) (a) Liebe, B.; Kunz, H. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 618. (b) Liebe, B.; Kunz, H. *Helv. Chim. Acta* **1997**, *80*, 1473. (c) Dziadek, S.; Espínola, C. G.; Kunz, H. *Aust. J. Chem.* **2003**, *56*, 519.

(8) George, S. K.; Holm, B.; Reis, C. A.; Schwientek, T.; Clausen, H.; Kihlberg, J. *J. Chem. Soc., Perkin Trans. 1* **2001**, 880.

(9) Helferich, B.; Wedemeyer, F. *Liebigs Ann. Chem.* **1949**, *563*, 139.

(10) Lergenmüller, M.; Ito, Y.; Ogawa, T. *Tetrahedron* **1988**, *54*, 1381.

(11) (a) Marra, A.; Sinaÿ, P. *Carbohydr. Res.* **1989**, *187*, 35. (b) Keil, S.; Claus, C.; Dippold, W.; Kunz, H. *Angew. Chem. Int. Ed.* **2001**, *40*, 366.

(12) Martichonok, V.; Whitesides, G. M. *J. Org. Chem.* **1996**, *61*, 1702.

(13) **9**: $[\alpha]_D^{27}$ = 32.01 (c = 1, $CHCl_3$); $R_t$ = 30.07 min [Phenomenex Luna C18 (2), grad.: $CH_3CN/H_2O$ (60:40) → (90:10), 40min]; $^1H$ NMR (400 MHz, $CDCl_3$, COSY, HMQC): δ (ppm) = 7.74 (d, 2 H, H4-, H5-Fmoc, $J_{H4,H3}$ = $J_{H5,H6}$ = 7.84 Hz), 7.58 (dd, 2 H, H1-, H8-Fmoc, $J_{H1,H2}$ = $J_{H8,H7}$ = 7.80 Hz), 7.49 (d, 2 H, $H_{ar}$-Bzn, $J_{Ha,Hb}$ = 5.84), 7.42–7.35 (m, 2 H, H3-, H6-Fmoc), 7.35–7.24 (m, 15 H, H2-, H7-Fmoc, $H_{ar}$-Bzn (3 H), $H_{ar}$-Bn (10 H)), 6.59 (d, 1 H, NH-GalNAc, $J_{NH,H2}$ = 9.40 Hz), 6.12 (d, 1 H, NH-T, $J_{NH,T\alpha}$ = 9.76 Hz), 5.45 (s, 1 H, CH-Bzn), 5.40 (t, 1 H, H8″, $J_{H8'',H9''}$ = 6.64 Hz), 5.28–5.13 (m, 2 H, H7″, NH-NeuNAc), 5.16 (s, 2 H, $CH_2$-Bn), 4.98 (d, 1 H, H1, $J_{H1,H2}$ = 3.52 Hz), 4.92–4.81 (m, 1 H, H4″), 4.74–4.63 (m, 1 H, H2), 4.55 (s, 2 H, $CH_2$-Bn), 4.52–4.03 (m, 10 H, $H9_a$″{4.46}, $CH_2$-Fmoc {4.39, 4.34}, $T^\beta$ {4.33}, H4 {4.26}, $T^\alpha$ {4.25}, H9-Fmoc {4.19}, H1′ {4.15}, $H6_a$ {4.13}, H5″ {4.06}), 3.99–3.84 (m, 4H, H6″ {3.93}, $H9_b$″ {3.92}, H3′ {3.90}, $H6_b$ {3.87}), 3.72–3.55 (m, 4 H, H3 {3.67}, H2′ {3.61}, $H6_a$′ {3.59}, H5 {3.57}), 3.54–3.39 (m, 2 H, $H6_b$′{3.87}, H5′{3.41}), 3.24 (s, 1 H, H4′), 2,78 ($s_b$, 1 H, OH), 2.71 (dd, 1 H, $H3_{eq}$″, $J_{H3eq'',H3ax''}$ = 12.72 Hz, $J_{H3eq'',H4''}$ = 4.28 Hz), 2.36 (s, 1 H, OH), 2.03–2.00 (m, 1 H, $H3_{ax}$″), 2.09, 2.07, 2.02, 2.00, 1.92, 1.82 (5*s, 18H, 6* $CH_3$-Ac), 1.42 (s, 9 H, $CH_3$-*t*-Bu), 1.26 (d, 3 H, $T^\gamma$, $J_{T\gamma,T\beta}$ = 5.92 Hz); $^{13}C$ NMR (100.6 MHz, $CDCl_3$, BB, HMQC): δ = 171.53, 170.91, 170.82, 170.24, 170.09, 169.76, 167.77 (C=O), 156.83 (C=O-urethane), 143.71 (C1a-, C8a-Fmoc), 141.29, 141.15 (C4a-, C5a-Fmoc), 138.07, 137.51 ($C_q$-Bn), 134.21 ($C_q$-Bn), 128.97, 128.84, 128.78, 128.43, 127.51 ($C_{ar}$-Bn), 128.78, 128.04, 126.44 ($C_{ar}$-Bzn), 127.79 (C3-, C6-Fmoc), 127.09, 126.97 (C2-, C7-Fmoc), 125.17, 124.94 (C1-, C8-Fmoc), 120.03 (C4-, C5-Fmoc), 106.55 (C1′), 100.96 (CH-Bzn), 100.45 (C1), 97.47 (C2″), 83.13 ($C_q$-*t*-Bu), 78.15 (C3), 75.94, 75.15 ($T^\beta$, C4, C3′), 73.60 ($CH_2$-Bn), 73.43, 72.92 (C6″,C5′), 69.72 (C8″, C6′), 69.07 (C6), 68.61 (C4″), 68.22, 68.11 ($CH_2$-Bn, C2′,C4′,C7″), 67.25 ($CH_2$-Fmoc), 63.64 (C9″), 63.52 (C5), 59.24 ($T^\alpha$), 49.08 (C5″), 47.79 (C2), 47.09 (C9-Fmoc), 37.42 (C3″), 28.06 ($CH_3$-*t*-Bu), 23.16, 23.07 ($CH_3$-NHAc), 21.33, 21.21, 21.01, 20.78, 20.64 ($CH_3$-OAc), 19.47 ($T^\gamma$), HR-ESI-TOF-MS (positive ion mode): calcd. for $C_{77}H_{91}N_3O_{27}Na$: 1512.5737, found: 1512.5729 $[M + Na]^+$.

(14) Neises, B.; Steglich, W. *Angew. Chem., Int. Ed. Engl.* **1978**, *17*, 522.

(15) **12**: $[\alpha]_D^{27}$ = 24.75 (c = 1, $CHCl_3$), HR-ESI-TOF-MS (positive ion mode): calcd. for $C_{74}H_{86}N_3O_{31}Na_2$: 1558.5040, found: 1558.5024 $[M - H + 2 \times Na]^+$.

(16) Perkin Elmer ABI 433A (Applied Biosystems).

(17) (a) Wagner, M.; Kunz, H. *Angew. Chem. Int. Ed.* **2002**, *41*, 317. (b) Wagner, M.; Kunz, H. *Z. Naturforsch.* **2002**, *57b*, 928.

(18) Knorr, R.; Trzeciak, A.; Bannwarth, W.; Gillessen, D. *Tetrahedron Lett.* **1989**, *30*, 1927.

(19) (a) Carpino, L. A. *J. Am. Chem. Soc.* **1993**, *115*, 4397. (b) Carpino, L. A.; El-Faham, A. S.; Minor, C.; Albericio, F. *J. Chem. Soc., Chem. Commun.* **1994**, 210.

(20) **16**: $[\alpha]_D^{26}$ = –40.68 (c = 0.86, $H_2O$), $R_t$ = 14.52 min [Phenomenex Jupiter C18, grad.: $CH_3CN/H_2O$ + 0.1% TFA (5:95) → (30:70), 40min]; $^1H$ NMR (400 MHz, $D_2O$, COSY, TOCSY, HMQC): δ (ppm) = 4.92 (d, 1 H, H1, $J_{H1,H2}$ = 3.72 Hz), 4.69 (t, 1 H, $D^\alpha$, $J_{D\alpha,D\beta}$ = 6.36 Hz), 4.65–4.58 (m, 2 H, $T_{ST}^\alpha$ {4.63}, $R^\alpha$ {4.62}), 4.57–4.49 (m, 1 H, $A_1^\alpha$), 4.49–4.41 (m, 3 H, H1′ {4.47}, $S^\alpha$ {4.45}, $A_2^\alpha$ {4.44}), 4.40–4.34 (m, 3 H, 3 × $P^\alpha$), 4.33–4.26 (m, 3 H, $T_{ST}^\beta$ {4.31}, $V^\alpha$ {4.29}, $T_1^\alpha$ {4.28}), 4.22–4.14 (m, 3 H, H4 {4.21}, H2 {4.19}, $T_1^\beta$ {4.18}), 4.08–3.95 (m, 3 H, H3′ {4.04}, H7″ {4.03}, H3 {4.00}), 3.94–3.87 (m, 3 H, $G^\alpha$ {3.91}, H4′{3.90}), 3.86–3.52 (m, 20 H, H5′, H8″ {3.83}, $H6_a$ {3.80}, $S^\beta$ {3.76}, 3 × $P_a^\delta$ {3.73}, H5″ {3.71}, $H9_a$″, $H9_b$″, H4″ {3.67}, 3 × $P_a^\alpha$ {3.63}, $H6_b$ {3.62}, H6″, H5, $H6_a$′, $H6_b$′), 3.48 (dd, 1 H, H2′, $J_{H1',H2'}$ = 7.92 Hz), 3.21–3.14 (m, 2 H, $R^\delta$), 2.99–2.83 (m, 2 H, $D^\beta$), 2.71 (dd, 1 H, $H3_{eq}$″, $J_{H3eq'',H3ax''}$ = 12.20 Hz, $J_{H3eq'',H4''}$ = 4.40 Hz), 2.35–2.19 (m, 3 H, 3 × $P_a^\beta$), 2.10–1.94 (m, 8 H, $V^\beta$ {2.03}, 3 × $P^\gamma$ {2.01}, $R_a^\beta$), 2.02, 1.99, 1.97 (3 × s, 9 H, 3 × $CH_3$-NHAc), 1.93–1.77 (m, 4 H, 3 × $P_b^\beta$ {1.88}, $H3_{ax}$″ {1.80}), 1.77–1.59 (m, 3 H, $R_b^\beta$ {1.70}, $R^\gamma$ {1.63}), 1.38–1.31 (m, 6 H, 2 × $A^\beta$), 1.24 (d, 3 H, $T_{ST}^\gamma$, $J_{T\gamma,T\beta}$ = 6.12 Hz), 1.15 (d, 3 H, $T_1^\gamma$, $J_{T\gamma,T\beta}$ = 6.36 Hz), 0.94 (d, 6 H, $V^\gamma$); MALDI-TOF-MS (dhb, positive ion mode): calcd. for $C_{76}H_{124}N_{17}O_{37}$: 1867.9, found: 1868.8 $[M + H]^+$, 1890.9 ($[M + Na]^+$, calcd.: 1889.9), 1907.1 ($[M + K]^+$, calcd.: 1905.9), 1912.8 ($[M - H + 2 \times Na]^+$, calcd.: 1911.9); ESI-MS (positive ion mode): calcd. for $[C_{76}H_{121}N_{17}O_{37}Na_4]^{2+}$: 977.88, found: 977.87 $[M - 2 H + 4 Na]^{2+}$.