Chemoenzymatic synthesis of derivatives of a T-cell-stimulating peptide which carry tumor-associated carbohydrate antigens

Shaji K. George,[a] Björn Holm,[a] Celso A. Reis,[bc] Tilo Schwientek,[b] Henrik Clausen[b] and Jan Kihlberg*[a]

[a] *Organic Chemistry, Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden*
[b] *School of Dentistry, University of Copenhagen, Nörre Allé 20, DK-2200 Copenhagen N, Denmark*
[c] *Institute of Molecular Pathology and Immunology of the University of Porto, IPATIMUP, Rua Dr. Roberto Frias s/n, 4200 Porto, Portugal*

Received (in Cambridge, UK) 27th November 2000, Accepted 1st February 2001
First published as an Advance Article on the web 2nd March 2001

The Tn (GalNAcα-Ser/Thr), T [Galβ(1→3)GalNAcα-Ser/Thr], sialyl-Tn [Neu5Acα(2→6)GalNAcα-Ser/Thr] and 2,3-sialyl-T [Neu5Acα(2→3)Galβ(1→3)GalNAcα-Ser/Thr] antigens are examples of tumor-associated carbohydrate antigens expressed by epithelial cancers. We now describe the preparation of 2-bromoethyl glycosides corresponding to the Tn and T antigens in one and five chemical steps (51 and 15% total yield), respectively, starting from *N*-acetylgalactosamine. The 2-bromoethyl Tn and T glycosides were used to alkylate a homocysteine residue incorporated in a peptide that is able to bind to class I MHC molecules on antigen-presenting cells. The two neoglycopeptides were then converted into glycopeptides which carry the sialyl-Tn and 2,3-sialyl-T antigens by using recombinant sialyltransferases. Interestingly, the sialyltransferases were able to sialylate the Tn and T carbohydrate moieties even though they were linked to the peptide backbone *via* a spacer instead of being attached to serine or threonine. The four glycopeptides will be used in studies directed towards inducing a carbohydrate-specific T cell response against the Tn, T, sialyl-Tn, and 2,3-sialyl-T antigens.

Introduction

Most epithelial cells produce mucins, which are a family of glycoproteins that are either secreted or exposed to the exterior in membrane-associated form.[1–3] Mucins have a rod-like appearance and are composed of a polypeptide backbone consisting of highly conserved tandem repeats, in which multiple serine and threonine residues carry complex carbohydrates. In epithelial cancers, such as those affecting the breast, ovary, stomach, and colon, mucins express simpler *O*-linked carbohydrates which are regarded as tumor-associated carbohydrate antigens.[4–8] The Tn (GalNAcα-Ser/Thr), T [Galβ(1→3)-GalNAcα-Ser/Thr], sialyl-Tn [Neu5Acα(2→6)GalNAcα-Ser/Thr] and 2,3-sialyl-T [Neu5Acα(2→3)Galβ(1→3)GalNAcα-Ser/Thr] antigens are important examples of such antigens.

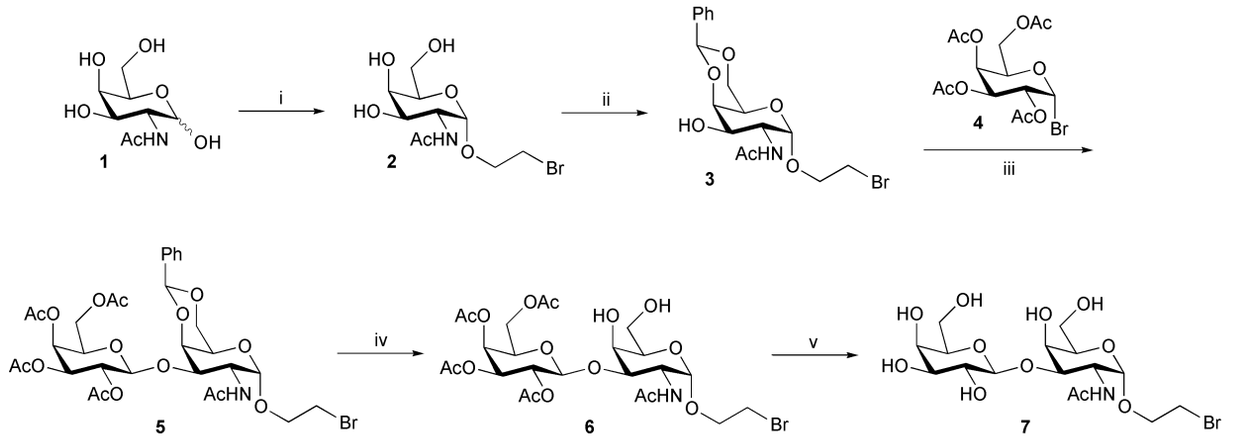
In view of their importance as tumor-associated antigens substantial efforts have been focused on chemical synthesis of glycopeptides containing the Tn, T, sialyl-Tn and 2,3-sialyl-T antigens. As for all *O*-linked glycopeptides the most general synthetic route in use today employs N^{α}-Fmoc-protected glycosylated amino acids as building blocks in stepwise assembly, preferably on solid phase.[9–12] Synthesis of building blocks corresponding to the structurally less complex Tn and T antigens has been described by several groups during recent years (reviewed in ref. 13). However, incorporation of sialic acid in glycoconjugates by chemical means is a demanding task,[14,15] which may explain why only a few syntheses of building blocks corresponding to the sialyl-Tn[16–19] and 2,3-sialyl-T[19–21] antigens have been reported. Recently, a few enzymatic syntheses of the 2,3-sialyl-T antigen have been described,[22,23] but since the carboxy group of the sialic acid residue in the resulting building blocks is unprotected they cannot be used for synthesis of glycopeptides.

Synthetic vaccines based on the tumor-associated Tn or sialyl-Tn antigens have been developed during recent years.[7,24,25] In these vaccines the antigens were coupled to an immunological carrier protein such as keyhole limpet hemocyanin (KLH) or to an immunogenic synthetic lipopeptide, either as individual glycosylated amino acids or after incorporation in short peptides. Studies in mice revealed that short Tn-based glycopeptides coupled to KLH induced a strong IgM response and a moderate IgG response, both of which were reactive to colon cancer cells.[25] Moreover, a sialyl-Tn KLH conjugate vaccine used in combination with cyclophosphamide was indicated to have a therapeutic affect when evaluated in clinical trials involving breast cancer patients.[26,27] These vaccines elicit an antibody response, but it would also be an advantage if T cells could be directed to tumor-associated carbohydrate antigens. T cells are known to recognize short peptides presented by major histocompatibility complex (MHC) molecules, but recent studies have also shown that glycopeptides can be recognized in a carbohydrate-specific manner.[11,28] Interestingly, one study performed in mice revealed that immunization with neoglycopeptides elicited a set of γ/δ cytotoxic T cells that were able to kill target cells expressing the same carbohydrate moiety in glycolipid form.[29] This study was based on the model carbohydrate galabiose [Galα(1→4)Galβ] which was linked to peptides known to be bound by class I MHC molecules. To allow investigations of whether it is possible to use this approach to direct T cells to the tumor-associated Tn, T, sialyl-Tn and 2,3-sialyl-T antigens we have now prepared 2-bromoethyl glycosides corresponding to the Tn and T antigens. These building blocks were then used in chemoenzymatic synthesis of four glycosylated derivatives of a peptide from vesicular stomatitis virus nucleocapsid protein, which binds to class I MHC molecules.

Results and discussion

Coupling of 3,4,6-tri-*O*-acetyl-2-azido-2-deoxy-α-D-galactopyranosyl bromide[30,31] to 2-bromoethanol under promotion by silver perchlorate and silver carbonate was investigated in an initial attempt to obtain an α-linked 2-bromoethyl glycoside of

 DOI: 10.1039/b009567m

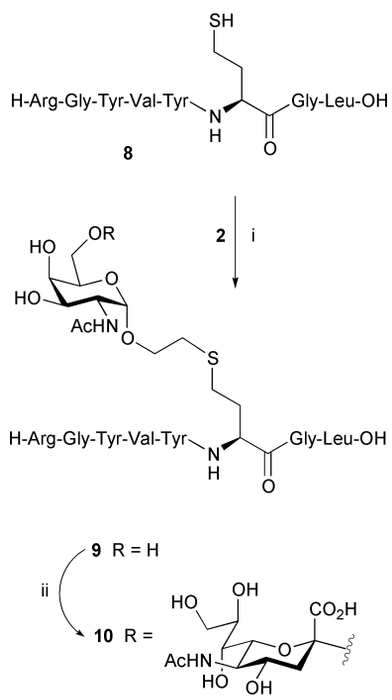


Scheme 1 *Reagents and conditions* (*and yields*): i) 2-Bromoethanol, HCl (g), 55–60 °C (52%); ii) α,α-dimethoxytoluene, *p*-TsOH (cat.), CH_3CN, rt (98%); iii) AgOTf, CH_2Cl_2–toluene (1 : 1), −25 °C (50%); iv) 80% aq. TFA, 0 °C (70%); v) NaOMe in MeOH, rt (84%).

N-acetylgalactosamine, *i.e.* an analogue of the Tn antigen. This approach was not successful since a mixture of the desired α-glycoside and the corresponding β-glycoside that could not be separated by flash column chromatography was obtained. Instead, it was found that 2-bromoethyl glycoside **2** could be synthesized from *N*-acetyl-D-galactosamine **1** in a Fischer glycosidation (Scheme 1). This was achieved by heating of **1** in 2-bromoethanol at 60 °C in the presence of a catalytic amount of hydrogen chloride to give **2** in 52% yield.[32] By performing the reaction under thermodynamic control a very short route to the desired α-glycoside **2** was thus found, in spite of the presence of a participating acetamido group at C-2 of *N*-acetylglucosamine.

Conversion of **2** into an analogue of the T antigen requires regioselective β-galactosylation of HO-3 (Scheme 1). In order to facilitate this HO-4 and HO-6 in **2** were first protected by treatment with α,α-dimethoxytoluene to afford **3** (98%). Formation of the β-galactosidic bond in derivatives of the T antigen is usually performed by employing derivatives of *N*-acetylgalactosamine as glycosyl acceptors, in which the acetamido group is masked as an azido group.[13] As compared with the acetamido group, the azido group reduces the steric hindrance for glycosylation at the adjacent HO-3 and also increases the nucleophilicity of this hydroxy group by avoiding an unfavourable hydrogen-bonding pattern.[33,34] After investigation of different glycosyl donors and promoters for glycoside synthesis it was found that glycosylation of **3** could be carried out directly using per-acetylated galactosyl bromide **4** to give the β-linked disaccharide **5** in 50% yield. Cleavage of the 4,6-benzylidene group in **5** was then achieved by using an ice-cold solution of aq. trifluoroacetic acid (TFA) to give **6** (70%). Finally, the acetyl groups in **6** were removed by treatment with methanolic sodium methoxide to give fully unprotected T antigen analog **7** (84%), after purification by reversed-phase HPLC.

Peptide **8** is an analog of an immunodominant, class I MHC restricted epitope from vesicular stomatitis virus nucleocapsid protein,[35] in which glutamine at position 6 has been replaced by homocysteine (Schemes 2 and 3). Compound **8** was prepared by Fmoc solid-phase synthesis using conditions described previously,[36] and was then purified by reversed-phase HPLC. Alkylation of the sulfhydryl group in **8** was accomplished by reaction with the unprotected 2-bromoethyl glycosides **2** and **7** using caesium carbonate as base.[37] The reaction was carried out in dry *N*,*N*-dimethylformamide (DMF) under an inert atmosphere of nitrogen in order to avoid oxidative dimerization of **8**. Alkylation of **8** was monitored by analytical reversed-phase HPLC and was generally found to be complete within an hour. The reaction was then quenched by addition of aq. TFA. Freeze-drying and purification by preparative reversed-phase HPLC gave neoglycopeptides **9** and **11** (74 and 58% yield, respectively), which carry the carbohydrate moieties of the Tn and T antigens. The structures of glycopeptides **9** and **11** were verified by amino acid analysis, mass spectrometry and ^1H NMR spectroscopy (Tables 1 and 2).



Scheme 2 *Reagents and conditions* (*and yields*): i) Cs_2CO_3, DMF, rt (74%); ii) CMP-Neu5Ac, α-2,6 sialyltransferase (ST6-GalNAc-I), Bis-Tris buffer (20 mM; pH 6.0), 37 °C (≈40%).

Recombinant mouse *N*-acetylgalactosamine α2-6 sialyltransferase[38] (ST6-GalNAc-I) was expressed in insect cells and purified by ion-exchange chromatography.[39] ST6-GalNAc-I was then used to transform **9** into **10**, which carries the carbohydrate moiety of the sialyl-Tn antigen (Scheme 2). This was achieved by incubating the enzyme with **9** and CMP-Neu5Ac (CMP = cytosine monophosphate) in Bis-Tris buffer (pH 6.0) at 37 °C. After 6 h, analysis[40] by nano-scale reversed-phase HPLC in combination with MALDI-TOF mass spectrometry indicated 90–95% conversion of **9** into **10**. When the sialylation was performed on a semipreparative scale (0.5 mg of **9**) glycopeptide **10** was obtained in ≈40% yield after purification by reversed-phase HPLC and freeze-drying. This result reveals that ST6-GalNAc-I can act on substrates in which the *N*-acetylgalactosamine residue being sialylated is linked to a different entity than serine or threonine.

Human core 1-specific α2-3 sialyltransferase[41] (ST3-Gal-I) was obtained by expression in Sf9 cells followed by purification

Table 1 ^{1}H NMR chemical shifts for glycopeptide **9**[a]

Residue	NH	H^{α}	H^{β}	H^{γ}	Others
Arg		3.99	1.87[b]	1.61[b]	3.15 (H-δ, -δ′), 7.20, 6.90 and 6.45 (NH)
Gly	8.78	4.00, 3.90			
Tyr	8.42	4.46	2.92, 2.84		7.01 and 6.75 (ArH)
Val	7.96	3.91	1.82	0.81, 0.76	
Tyr	8.48	4.37	3.02, 2.86		7.15 and 6.75 (ArH)
Hcy[c]	8.37	4.36	1.80, 2.02	2.52, 2.38	2.70[b] (SCH_2CH_2O), 3.78 and 3.56 (SCH_2CH_2O)
Gly	7.00	3.76[b]			
Leu	7.90	4.16	1.58[b]	1.55	0.87 and 0.83 (γ-, γ′-Me)

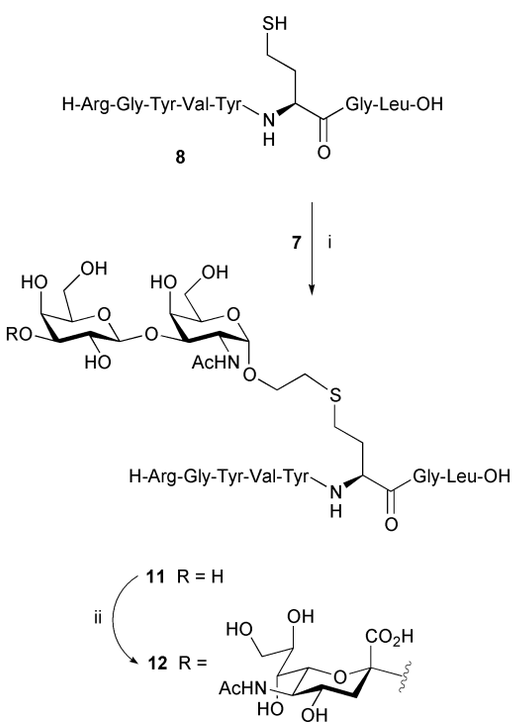
[a] Spectra were recorded at 500 MHz for a solution in H_2O–D_2O (9 : 1; pH 5.8) at 278 K using water (δ_H 4.98) as internal standard. [b] Degeneracy has been assumed. [c] Chemical shifts (δ) for the GalNAc moiety in **9**: 8.14 (NH), 4.85 (H-1), 4.09 (H-2), 3.84 (H-3), 1.96 ($NHCOCH_3$).

Table 2 ^{1}H NMR chemical shifts for glycopeptide **11**[a]

Residue	NH	H^{α}	H^{β}	H^{γ}	Others
Arg		3.93	1.81[b]	1.56[b]	3.11[b] (H-δ, -δ′), 7.20 (CH_2NH)
Gly	8.74	3.97, 3.87			
Tyr	8.38	4.46	2.90, 2.81		6.99 and 6.77 (ArH)
Val	7.93	3.92	1.79	0.76[b]	
Tyr	8.45	4.37	2.99, 2.84		7.12 and 6.74 (ArH)
Hcy[c]	8.34	4.38	1.99, 1.78	2.49, 2.34	2.73[b] (SCH_2CH_2O), 3.79 and 3.59 (SCH_2CH_2O)
Gly	7.08	3.76[b]			
Leu	7.86	4.16	1.53[b]	1.53	0.83[b] (γ-, γ′-Me)

[a] Spectra were recorded at 500 MHz for a solution in H_2O–D_2O (9 : 1; pH 5.8) at 278 K using water (δ_H 4.98) as internal standard. [b] Degeneracy has been assumed. [c] Chemical shifts (δ) for the Galβ(1→3)GalNAc moiety in **11**: 8.19 (NH), 4.81 (H-1), 4.37 (H-1′), 4.27 (H-2), 3.97 (H-3), 3.85 (H-4′), 3.67 (H-4), 3.58 (H-5), 3.56 (H-3′), 3.47 (H-2′), 1.96 ($NHCOCH_3$).

by ion-exchange and gel-filtration chromatography.[23] Purified ST3-Gal-I was used for sialylation of neoglycopeptide **11** to give **12**, which carries the 2,3-sialyl-T antigen (Scheme 3). The



Scheme 3 *Reagents and conditions* (*and yields*): i) Cs_2CO_3, DMF, rt (58%); ii) CMP-Neu5Ac, α-2,3 sialyltransferase (ST3-Gal-I), calf intestinal phosphatase, Tris-HCl buffer (25 mM; pH 6.5), 37 °C (68%).

sialylation was performed at 37 °C in Tris-HCl buffer (pH 6.5) containing CMP-Neu5Ac as sialic acid donor. Calf intestinal phosphatase was added in order to prevent product inhibition of the sialyltransferase by CMP. After 3 h, nano-scale reversed-phase HPLC in combination with MALDI-TOF mass spectrometry[40] indicated ≈95% conversion of **11** into **12**. By performing repeated sialylations on a semipreparative scale (7 × 0.5 mg of **11**) glycopeptide **12** could be obtained in ≈70% yield after purification by HPLC and freeze-drying. Assignment of the ^{1}H NMR spectrum of **12** (Table 3) revealed NOEs between H-3 of the galactose moiety and both H-3^{ax} and H-3^{eq} of the sialic acid residue, thereby confirming that sialylation had occurred on HO-3 of the galactose moiety as expected. A side-product (≈20%) with a mass 16 daltons higher than that of **12** was also obtained in the sialylation. This was believed to originate from oxidation of the sulfur atom in the linker between the carbohydrate and peptide moieties of **12**, a suggestion that was further supported by ^{1}H NMR chemical-shift differences for the resonances in the linker of the side-product as compared with glycopeptide **12**.

Conclusions

The Tn, T, sialyl-Tn and 2,3-sialyl-T antigens are important tumor-associated antigens in epithelial cancers affecting the breast, ovary, stomach, and colon.[4–8] Herein we have described the preparation of 2-bromoethyl glycosides **2** and **7**, which correspond to the Tn and T antigens, in one and five chemical steps, respectively. The 2-bromoethyl glycosides were then converted into neoglycopeptides **9** and **11** by chemical methods. Finally, enzymatic sialylation of **9** and **11** gave glycopeptides **10** and **12**, which carry carbohydrate moieties corresponding to the sialyl-Tn and 2,3-sialyl-T antigens. Chemical incorporation of sialic acid in glycoconjugates is generally regarded as a demanding task.[14,15] This is supported by the fact that only a few, multistep syntheses of sialyl-Tn[16–19] and 2,3-sialyl-T[19–21] antigen building blocks using organochemical methods have been reported in the literature. Moreover, only in one case was such a 2,3-sialyl-T building block incorporated in a glycopeptide.[20] As demonstrated in the present paper, chemical synthesis of glycopeptides that carry the simpler Tn and T antigens, followed by enzymatic sialylation to give the more complex sialyl-Tn and 2,3-sialyl-T antigens, constitutes a more efficient route to these glycopeptides. It is interesting to note that the sialyltransferases employed in this study were able to sialylate the Tn and T antigens even though they were linked to the peptide backbone

Table 3 ^{1}H NMR chemical shifts for glycopeptide **12**[a]

Residue	NH	H^{α}	H^{β}	H^{γ}	Others
Arg		3.99	1.88[b]	1.62[b]	3.16[b] (H-δ, -δ′), 7.19 (CH_2NH)
Gly	8.78	4.02, 3.90			
Tyr	8.43	4.48	2.94, 2.85		7.03 and 6.74 (ArH)
Val	7.97	3.94	1.83	0.82, 0.78	
Tyr	8.49	4.40	3.03, 2.87		7.16 and 6.78 (ArH)
Hcy[c]	8.37	4.42	2.04, 1.81	2.52, 2.36	2.72[b] (SCH_2CH_2O), 3.80 and 3.60 (SCH_2CH_2O)
Gly	7.01	3.80[b]			
Leu	7.90	4.19	1.57[b]	1.57	0.88 and 0.84 (γ-, γ′-Me)

[a] Spectra were recorded at 500 MHz for a solution in H_2O–D_2O (9 : 1; pH 5.4) at 278 K using water (δ_H 4.98) as internal standard. [b] Degeneracy has been assumed. [c] Chemical shifts (δ) for the Neu5Acα(1→3)Galβ(1→3)GalNAc moiety in **12**. Neu5Ac: 8.15 (NH), 3.82 (H-5), 3.64 (H-4), 3.57 (H-6), 2.71 (H-3^{eq}), 2.00 ($NHCOCH_3$), 1.76 (H-3^{ax}). Gal: 4.47 (H-1), 4.03 (H-3), 3.50 (H-2). GalNAc: 8.21 (NH), 4.87 (H-1), 4.27 (H-2), 4.19 (H-4), 3.98 (H-3), 1.98 (Ac).

via a spacer, instead of being attached to serine or threonine. Chemoenzymatic synthesis thus appears to be able to provide access to both natural glycopeptides, *i.e.* those having the carbohydrate moieties linked to serine or threonine, and to neoglycopeptides which contain the tumor-associated sialyl-Tn and 2,3-sialyl-T antigens.

Experimental

General methods and materials

All reactions were carried out under a nitrogen atmosphere and anhydrous conditions using dry solvents unless otherwise stated. CH_2Cl_2 was dried by distillation from calcium hydride whereas THF and toluene were distilled from sodium benzophenone ketyl. Pyridine was dried over 4 Å molecular sieves. DMF was distilled and then dried over flame-dried 3 Å molecular sieves. TLC was performed on silica gel 60 F_{254} (Merck) with detection by UV light, or by charring with aq. sulfuric acid (10%). Flash column chromatography was performed on silica gel (Matrex, 60 Å; 35–70 μm, Grace Amicon) with distilled solvents. Organic solutions were dried over anhydrous Na_2SO_4 before being concentrated.

The ^{1}H and ^{13}C NMR spectra for compounds **2**, **3**, and **5**–**7** were recorded at 298 K with a Bruker DRX-400 spectrometer at 400 MHz and 100 MHz, respectively. NMR samples of these compounds were dissolved in $CDCl_3$ [residual $CHCl_3$ (δ_H 7.26) or $CDCl_3$ (δ_C 77.0) as internal standard] or CD_3OD [residual CD_2HOD (δ_H 3.35), CD_3OD (δ_C 49.0) as internal standard]. ^{1}H NMR spectra of glycopeptides **9**, **11** and **12** were recorded with a Bruker ARX-500 spectrometer for solutions in a 9 : 1 mixture of water and D_2O [H_2O (δ_H 4.98) as internal standard] at 278 K. Proton resonances for **9**, **11** and **12** were assigned from COSY, NOESY and TOCSY experiments. Optical rotations were recorded on a Perkin-Elmer 343 polarimeter and $[\alpha]_D$-values are given in 10^{-1} deg cm^2 g^{-1}. Analytical reversed-phase HPLC was performed on a Kromasil C-8 column (250 × 4.6 mm; 5 μm; 100 Å), eluted with a linear gradient of MeCN in water containing 0.1% TFA (flow-rate of 1.5 mL min^{-1}; detection at 214 nm). Preparative reversed-phase HPLC was performed on a Kromasil C-8 column (250 × 20 mm; 5 μm; 100 Å) using the same solvent system (flow-rate of 11 mL min^{-1}; detection at 214 nm).

2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl bromide **4** was prepared from per-acetylated galactose by treatment with HBr in HOAc–Ac_2O at 0 °C. L-Arginylglycyl-L-tyrosyl-L-valyl-L-tyrosyl-L-homocysteinylglycyl-L-leucine[37] **8** was prepared manually by Fmoc solid-phase synthesis in a mechanically agitated reactor using conditions described previously.[36] Recombinant mouse *N*-acetylgalactosamine α2–6 sialyltransferase[38] (ST6-GalNAc-I) was expressed in insect cells. The enzyme was purified by sequential ion-exchange chromatography on Amberlite (IRA95, Sigma) and SP-sepharose (Sigma) essentially as described.[39] Purified samples were concentrated in a centrifugation cartridge (Sigma) before use. Human core 1-specific α2–3 sialyltransferase[41] (ST3-Gal-I) was expressed and partially purified as described.[23] Final purification of ST3-Gal-I was performed on MiniS™ (PC3.2/3) using the Smart system (Pharmacia).

2-Bromoethyl 2-acetamido-2-deoxy-α-D-galactopyranoside 2

A mixture of *N*-acetyl-D-galactosamine **1** (Sigma, 1.50 g, 6.75 mmol) and 2-bromoethanol (20 g, 160 mmol) containing 2% of HCl (gas) was stirred at 55–60 °C for 5 h. The mixture was allowed to attain room temperature and the solvents were evaporated off. The residue was co-concentrated several times from toluene. Flash column chromatography of the residue (CH_2Cl_2–MeOH 65 : 20) gave **2** (1.14 g, 52%) as a white amorphous solid, $[\alpha]_D^{20}$ +141.8 (*c* 0.53, methanol); ^{1}H NMR (400 MHz; CD_3OD) δ 4.91 (1H, d, *J* 3.8 Hz, H-1), 4.17 (1H, dd, *J* 11, 3.7 Hz, H-2), 3.88 (1H, m, OCH_2CH_2Br), 3.80 (1H, m, H-5), 3.81–3.77 (1H, m, H-4), 3.72–3.65 (1H, m, H-3), 3.65–3.56 (1H, m, OCH_2CH_2Br), 3.65–3.56 (2H, m, H_2-6), 3.48 (2H, m, $BrCH_2$), 1.79 (3H, s, CH_3CO); ^{13}C NMR (100 MHz; CD_3OD) δ_C 172.5, 98.0, 71.5, 69.0, 68.2, 68.0, 61.2, 50.0, 30.2, 21.2; HRMS (FAB) Calc. for $C_{10}H_{18}BrNO_6$·Na: 350.0215 (M + Na$^+$). Found: *m*/*z*, 350.0211.

2-Bromoethyl 2-acetamido-4,6-*O*-benzylidene-2-deoxy-α-D-galactopyranoside 3

A solution of **2** (302 mg, 0.92 mmol), α,α-dimethoxytoluene (281 mg, 1.84 mmol) and a catalytic amount of toluene-*p*-sulfonic acid (*p*-TsOH) (17 mg) in acetonitrile (17 mL) was stirred at room temperature for 24 h. The reaction was quenched by addition of triethylamine (0.1 mL), the solvents were evaporated off, and the residue was dried under vacuum. Flash column chromatography of the residue (EtOAc–MeOH 96 : 4) afforded **3** (357 mg, 98%) as an amorphous white solid; $[\alpha]_D^{20}$ +114.3 (*c* 0.64, $CHCl_3$); ^{1}H NMR ($CDCl_3$) δ 7.53–7.35 (5H, m, ArH), 5.93 (1H, d, *J* 8.9 Hz, AcN*H*), 5.58 (1H, s, PhC*H*), 5.05 (1H, d, *J* 3.6 Hz, H-1), 4.48 (1H, ddd, *J* 10.9, 9.0, 3.6 Hz, H-2), 4.27 (1H, dd, *J* 12.6, 1.6 Hz, H-4), 4.24 (1H, m, H-3), 4.10–4.03 (2H, m, H-5 and OCH_2CH_2Br), 3.85–3.78 (3H, m, OCH_2CH_2Br and H_2-6), 3.56 (1H, t, *J* 5.1 Hz, CH_2Br), 3.55 (1H, t, *J* = 5.1 Hz, CH_2Br), 2.1 (3H, s, CH_3CO); ^{13}C NMR (400 MHz; $CDCl_3$) δ_C 171.2, 137.2, 129.0, 128.0, 126.1, 101.2, 98.7, 75.5, 69.4, 69.2, 68.0, 63.4, 50.2, 31.2, 23.5; HRMS (FAB) Calc. for $C_{17}H_{23}BrNO_6$: 416.0709 (M + H$^+$). Found: *m*/*z*, 416.0708.

2-Bromoethyl 2-acetamido-4,6-*O*-benzylidene-2-deoxy-3-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-α-D-galactopyranoside 5

A solution of silver triflate (465 mg, 1.81 mmol) in toluene (10 mL) was added dropwise to a suspension of **3** (500 mg, 1.2 mmol), 2,3,4,6-tetra-*O*-acetyl-α-D-galactopyranosyl bromide **4** (738 mg, 1.80 mmol) and crushed molecular sieves (flame-dried,

4 Å; 1.0 g) in CH_2Cl_2 (10 mL) at −25 °C. After 1 h, the reaction was quenched by addition of pyridine (1 mL) while keeping the temperature at −25 °C. After 5 min the mixture was allowed to attain room temperature, and the solids were removed by filtration (Hyflow, Supercel) and washed with CH_2Cl_2 (30 mL). The combined filtrates were washed with a mixture of aq. $Na_2S_2O_3$ (0.5 M) and saturated aq. $NaHCO_3$ (1 : 1; 20 mL), dried and concentrated. Repeated flash column chromatography of the residue (EtOAc–MeOH 100 : 0.5) afforded **5** as a white amorphous solid (447 mg, 50%), $[a]_D^{20}$ +94.4 (*c* 0.45, $CHCl_3$); ^{1}H NMR (400 MHz; $CDCl_3$) δ 7.52 and 7.50 (2H, 2d, ArH), 7.37–7.30 (3H, m, ArH), 5.76 (1H, d, *J* 8.8 Hz, NH), 5.55 (1H, s, PhC*H*), 5.36 (1H, dd, *J* 3.4, 1.1 Hz, H-4′), 5.17 (1H, dd, *J* 10.4, 7.9 Hz, H-2′), 5.15 (1H, d, *J* 3.6 Hz, H-1), 4.97 (1H, dd, *J* 10.4, 3.5 Hz, H-3′), 4.75 (1H, d, *J* 7.9 Hz, H-1′), 4.64 (1H, m, H-2), 4.29 (1H, m, H-4), 4.24–4.06 (5H, m, H_2-6, -6′ and H-5′), 4.03 (1H, m, OCH_2CH$_2$Br), 3.99 (1H, dd, *J* 5.9, 0.9 Hz, H-3), 3.80 (1H, m, OCH_2CH$_2$Br), 3.75 (1H, br s, H-5), 3.57 (2H, m, CH_2Br), 2.13, 2.05, 2.00, 1.97 and 1.94 (15H, 5s, CH_3CO); ^{13}C NMR ($CDCl_3$) δ_C 170.1, 166.5, 166.5, 165.2, 137.1, 133.5, 129.9, 129.8, 129.7, 129.6, 128.6, 128.5, 128.3, 128.3, 128.2, 127.8, 125.6, 100.2, 98.5, 93.9, 72.7, 71.7, 69.2, 69.0, 68.7, 68.3, 68.0, 66.8, 63.2, 62.4, 47.7, 31.8, 30.7, 29.0, 23.3, 22.6; HRMS (FAB) Calc. for $C_{31}H_{41}BrNO_{15}$: 746.1660 (M + H$^+$). Found: *m*/*z*, 746.1677.

2-Bromoethyl 2-acetamido-2-deoxy-3-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-α-D-galactopyranoside 6

An ice-cold solution of TFA in water (10 mL; 80%) was added during 5 min to a solution of **5** (204 mg, 0.27 mmol) in CH_2Cl_2 (5 mL) at 0 °C. After stirring of the mixture for 1 h, toluene (25 mL) was added and the mixture was concentrated. Flash column chromatography of the residue (EtOAc–MeOH 100 : 1) gave **6** (143 mg, 70%) as an amorphous white solid, $[a]_D^{20}$ +54.0 (*c* 0.62, $CHCl_3$); ^{1}H NMR ($CDCl_3$) δ 6.04 (1H, d, *J* 8.9 Hz, NH), 5.35 (1H, m, H-4′), 5.15 (1H, dd, *J* 10.2, 7.9 Hz, H-2′), 4.99 (1H, dd, *J* 10.5, 3.2 Hz, H-3′), 4.90 (1H, d, *J* 2.7 Hz, H-1), 4.66 (1H, d, *J* 7.7 Hz, H-1′), 4.54 (1H, m, H-2), 4.24–3.99 (4H, m, H-6, -6′, OCH_2CH$_2$Br and H-5), 3.99–3.87 (2H, m, H-6 and -6′), 3.87–3.72 (2H, m, OCH_2CH$_2$Br and H-3), 3.49 (2H, m, CH_2Br), 2.13, 2.05, 2.03, 1.99 and 1.95 (15H, 5s, 5 × CH_3CO); ^{13}C NMR ($CDCl_3$) δ_C 170.5, 170.2, 170.0, 169.8, 101.1, 97.9, 77.9, 70.8, 70.5, 69.7, 69.5, 68.7, 67.8, 66.9, 62.5, 61.3, 47.9, 31.6, 20.6, 20.5, 20.5, 20.4, 20.3; HRMS (FAB) Calc. for $C_{24}H_{36}BrNO_{15}$·Na: 680.1166 (M + Na$^+$). Found: *m*/*z*, 680.1161.

2-Bromoethyl 2-acetamido-2-deoxy-3-*O*-β-D-galactopyranosyl-α-D-galactopyranoside 7

A solution of **6** (140 mg, 0.286 mmol) in methanol (6 mL) was treated with methanolic NaOMe (0.2 M; 0.28 mL). After 1 h the solution was neutralized by addition of acetic acid in methanol (10%) and was then concentrated. The residue was purified using reversed-phase HPLC (gradient 0→100% CH_3CN in water, both containing 0.5% TFA, during 60 min; t_R 7.65 min) followed by freeze drying to give **7** (86 mg, 84%) as a white amorphous solid, $[a]_D^{20}$ +92.8 (*c* 0.53, MeOH); ^{1}H NMR (CD_3OD) δ 4.92 (1H, d, *J* 3.8 Hz, H-1), 4.43 (1H, dd, *J* 7.3, 3.7 Hz, H-2), 4.40 (1H, d, *J* 7.5 Hz, H-1′), 4.18 (1H, m, H-4), 4.00 (1H, m, OCH_2CH$_2$Br), 3.98–3.88 (2H, m, H-5 or -5′ and -3), 3.85–3.75 (2H, m, H-4′ and OCH_2CH$_2$Br), 3.75–3.68 (4H, m, H_2-6 and -6′), 3.62–3.55 (2H, m, CH_2Br), 3.55–3.49 (2H, m, H-2′ and -5′ or -5), 3.45 (1H, dd, *J* 9.7, 3.3 Hz, H-3′), 2.00 (3H, s, CH_3CO); ^{13}C NMR ($CDCl_3$) δ_C 174.3, 106.5, 99.4, 79.0, 76.9, 74.9, 72.7, 72.6, 70.4, 69.6, 63.0, 62.8, 50.4, 31.9, 23.0; HRMS (FAB) Calc. for $C_{16}H_{28}BrNO_{11}$·Na: 512.0743 (M + Na$^+$). Found: *m*/*z*, 512.0749.

Tn neoglycopeptide 9

Peptide **8** (15 mg, 13 μmol) was added to a mixture of 2-bromoethyl glycoside **2** (15 mg, 46 μmol) and caesium carbonate (50 mg, 154 μmol) in dry DMF (1.3 mL) under nitrogen. The solution was stirred at room temperature until analytical reversed-phase HPLC (gradient 0→100% CH_3CN in water, both containing 0.1% TFA, during 60 min; t_R 17.2 min) indicated that **8** was consumed. The reaction was quenched by addition of 0.1% aq. TFA (15 mL) and the mixture was freeze dried. Purification of the residue by preparative reversed-phase HPLC (gradient 0→100% CH_3CN in water, both containing 0.1% TFA, during 60 min) gave **9** (14 mg, 77% peptide content, 74% yield) as an amorphous white solid. ^{1}H NMR data, see Table 1; HRMS (FAB) Calc. for $C_{53}H_{83}N_{12}O_{17}S$: 1191.5720 (M + H$^+$). Found: *m*/*z*, 1191.5728; amino acid analysis: Gly 1.99 (2), Val 1.04 (1), Leu 1.01 (1), Tyr 1.98 (2), Arg 0.98 (1).

Sialyl-Tn neoglycopeptide 10

Neoglycopeptide **9** (0.50 mg, 0.42 μmol) was added to purified ST6-GalNAc-I (≈2 mU) in 20 mM Bis-Tris buffer (pH 6.0; 1.0 mL) containing CMP-Neu5Ac (2 mM), EDTA (20 mM) and dithiothreitol (1 mM). The solution was then incubated at 37 °C for 6 h, after which analysis[40] by nano-scale reversed-phase HPLC in combination with MALDI-TOF mass spectrometry indicated 90–95% conversion of **9**. Purification by HPLC on a Zorbax 300SB-C3 column (9.4 × 250 mm) using a gradient of 0→90% CH_3CN in water, both containing 0.1% TFA, gave **10** (250 μg, ≈40%) as an amorphous white solid after freeze-drying; MS (MALDI-TOF, linear mode, 2,5-dihydroxybenzoic acid) Calc. for $C_{64}H_{99}N_{13}O_{25}S$: 1482 (M$^+$). Found: *M*, 1483.

T neoglycopeptide 11

Peptide **8** (12 mg, 13 μmol) was added to a mixture of a 2-bromoethyl glycoside **7** (10 mg, 20 μmol) and caesium carbonate (35 mg, 107 μmol) in dry DMF (1.0 mL) under nitrogen. The mixture was stirred at room temperature until analytical reversed-phase HPLC (gradient 0→100% CH_3CN in water, both containing 0.1% TFA, during 60 min, t_R 16.9 min) indicated that **8** was consumed. The reaction was quenched by addition of 0.1% aq. TFA (15 mL) and the mixture was freeze dried. Purification of the residue by preparative reversed-phase HPLC (gradient 0→100% CH_3CN in water, both containing 0.1% TFA, during 60 min) gave **11** (10 mg, 81% peptide content, 58% yield) as an amorphous white solid. ^{1}H NMR data, see Table 2; HRMS (FAB) Calc. for $C_{59}H_{93}N_{12}O_{22}S$: 1353.6248 (M + H$^+$). Found: *m*/*z*, 1353.6215; amino acid analysis: Gly 1.99 (2), Val 0.99 (1), Leu 1.03 (1), Tyr 2.00 (2), Arg 1.00 (1).

2,3-Sialyl-T neoglycopeptide 12

Purified ST3-Gal-I (20 μL, ≈2 mU) was added to neoglycopeptide **11** (0.50 mg, 0.37 μmol) in 25 mM Tris-HCl buffer (pH 6.5; 1.0 mL) containing CMP-Neu5Ac (2 mM), Triton X-100 (0.1%) and calf intestinal phosphatase (10 mU). The solution was then incubated at 37 °C for 3 h, after which analysis[40] by nano-scale reversed-phase HPLC in combination with MALDI-TOF mass spectrometry indicated ≈95% conversion of **11**. This procedure was repeated so that 3.5 mg (2.6 μmol) of **11** was sialylated. Purification by HPLC on a Zorbax 300SB-C3 column (9.4 × 250 mm) using a gradient of 0→90% CH_3CN in water, both containing 0.1% TFA, gave **12** (2.9 mg, 68%) as an amorphous white solid after freeze-drying. ^{1}H NMR data, see Table 3; MS (ES) Calc. for $C_{70}H_{109}N_{13}O_{30}S$: 1643.7 (M$^+$). Found: *m*/*z*, 1643.6.

Acknowledgements

This article is dedicated (by J. K.) to the memory of Göran Magnusson, an outstanding scientist, supervisor and a wise mentor. This research was supported by a postdoctoral research fellowship to S. K. G. from the program 'Glycoconjugates

in Biological Systems (GLIBS)' sponsored by the Swedish Foundation for Strategic Research. Financial support from the EU Biotech 5th Framework and by FCT (Sapiens 36376/99) is also acknowledged.

References

1 K. L. Carraway and S. R. Hull, *Glycobiology*, 1991, **1**, 131.
2 I. Carlstedt and J. R. Davies, *Biochem. Soc. Trans.*, 1997, **25**, 214.
3 F.-G. Hanisch and S. Müller, *Glycobiology*, 2000, **10**, 439.
4 G. F. Springer, *Science*, 1984, **224**, 1198.
5 S.-i. Hakomori, *Adv. Cancer Res.*, 1989, **52**, 257.
6 S. H. Itzkowitz, M. Yuan, C. K. Montgomery, T. Kjeldsen, H. K. Takahashi, W. L. Bigbee and Y. S. Kim, *Cancer Res.*, 1989, **49**, 197.
7 J. Taylor-Papadimitriou, J. Burchell, D. W. Miles and M. Dalziel, *Biochim. Biophys. Acta*, 1999, **1455**, 301.
8 T. Irimura, K. Denda, S.-i. Iida, H. Takeuchi and K. Kato, *J. Biochem.* (*Tokyo*), 1999, **126**, 975.
9 M. Meldal, in *Glycopeptide Synthesis*, ed. Y. C. Lee and R. T. Lee, Academic Press, San Diego, 1994, p. 145.
10 T. Norberg, B. Lüning and J. Tejbrant, *Methods Enzymol.*, 1994, **247**, 87.
11 J. Kihlberg and M. Elofsson, *Curr. Med. Chem.*, 1997, **4**, 79.
12 G. Arsequell and G. Valencia, *Tetrahedron: Asymmetry*, 1997, **8**, 2839.
13 Y. Nakahara, H. Iijima and T. Ogawa, in *Stereocontrolled Approaches to O-Glycopeptide Synthesis*, ed. P. Kovác, American Chemical Society, Washington DC, 1994, p. 249.
14 O. Kanie and O. Hindsgaul, *Curr. Opin. Struct. Biol.*, 1992, **2**, 674.
15 M. Meldal and P. M. St Hilaire, *Curr. Opin. Chem. Biol.*, 1997, **1**, 552.
16 Y. Nakahara, H. Iijima, S. Shibayama and T. Ogawa, *Carbohydr. Res.*, 1991, **216**, 211.
17 B. Liebe and H. Kunz, *Angew. Chem.*, *Int. Ed. Engl.*, 1997, **36**, 618.
18 M. Elofsson, L. A. Salvador and J. Kihlberg, *Tetrahedron*, 1997, **53**, 369.
19 J. B. Schwarz, S. D. Kuduk, X.-T. Chen, D. Sames, P. W. Glunz and S. J. Danishefsky, *J. Am. Chem. Soc.*, 1999, **121**, 2662.
20 Y. Nakahara, Y. Nakahara, Y. Ito and T. Ogawa, *Carbohydr. Res.*, 1998, **309**, 287.
21 S. Komba, M. Meldal, O. Werdelin, T. Jensen and K. Bock, *J. Chem. Soc.*, *Perkin Trans. 1*, 1999, 415.
22 U. Gambert and J. Thiem, *Eur. J. Org. Chem.*, 1999, 107.
23 G. Dudziak, N. Bézay, T. Schwientek, H. Clausen, H. Kunz and A. Liese, *Tetrahedron*, 2000, **56**, 5865.
24 T. Toyokuni and A. K. Singhal, *Chem. Soc. Rev.*, 1995, **24**, 231.
25 S. J. Danishefsky and J. R. Allen, *Angew. Chem.*, *Int. Ed.*, 2000, **39**, 836.
26 G. D. MacLean, M. A. Reddish, R. R. Koganty and B. M. Longenecker, *J. Immunother.*, 1996, **19**, 59.
27 G. D. MacLean, D. W. Miles, R. D. Rubens, M. A. Reddish and B. M. Longenecker, *J. Immunother.*, 1996, **19**, 309.
28 F. R. Carbone and P. A. Gleeson, *Glycobiology*, 1997, **7**, 725.
29 U. M. Abdel-Motal, L. Berg, A. Rosén, M. Bengtsson, C. J. Thorpe, J. Kihlberg, J. Dahmén, G. Magnusson, K.-A. Karlsson and M. Jondal, *Eur. J. Immunol.*, 1996, **26**, 544.
30 R. U. Lemieux and R. M. Ratcliffe, *Can. J. Chem.*, 1979, **57**, 1244.
31 J. Broddefalk, U. Nilsson and J. Kihlberg, *J. Carbohydr. Chem.*, 1994, **13**, 129.
32 D. Qiu, S. S. Gandhi and R. R. Koganty, *Tetrahedron Lett.*, 1996, **37**, 595.
33 R. R. Schmidt and P. Zimmermann, *Angew. Chem.*, *Int. Ed. Engl.*, 1986, **25**, 725.
34 R. Polt, L. Szabó, J. Treiberg, Y. Li and V. J. Hruby, *J. Am. Chem. Soc.*, 1992, **114**, 10249.
35 D. H. Fremont, M. Matsumura, E. A. Stura, P. A. Peterson and I. A. Wilson, *Science*, 1992, **257**, 919.
36 B. Holm, J. Broddefalk, S. Flodell, E. Wellner and J. Kihlberg, *Tetrahedron*, 2000, **56**, 1579.
37 M. Bengtsson, J. Broddefalk, J. Dahmén, K. Henriksson, J. Kihlberg, H. Lönn, B. R. Srinivasa and K. Stenvall, *Glycoconjugate J.*, 1998, **15**, 223.
38 N. Kurosawa, S. Takashima, M. Kono, Y. Ikehara, M. Inoue, Y. Tachida, H. Narimatsu and S. Tsuji, *J. Biochem.* (*Tokyo*), 2000, **127**, 845.
39 H. H. Wandall, H. Hassan, E. Mirgorodskaya, A. K. Kristensen, P. Roepstorff, E. P. Bennett, P. A. Nielsen, M. A. Hollingsworth, J. Burchell, J. Taylor-Papadimitriou and H. Clausen, *J. Biol. Chem.*, 1997, **272**, 23503.
40 J. Gobom, E. Nordhoff, E. Mirgorodskaya, R. Ekman and P. Roepstorff, *J. Mass Spectrom.*, 1999, **34**, 105.
41 M. L. Chang, R. L. Eddy, T. B. Shows and J. T. Lau, *Glycobiology*, 1995, **5**, 319.