Formation of lactones from sialylated MUC1 glycopeptides

Maciej Pudelko,[a] Anna Lindgren,[a] Tobias Tengel,[b] Celso A. Reis,[c] Mikael Elofsson*[a] and Jan Kihlberg*[a,d]

Received 24th October 2005, Accepted 22nd December 2005
First published as an Advance Article on the web 19th January 2006
DOI: 10.1039/b514918e

The tumor-associated carbohydrate antigens T_N, T, sialyl T_N and sialyl T are expressed on mucins in several epithelial cancers. This has stimulated studies directed towards development of glycopeptide-based anticancer vaccines. Formation of intramolecular lactones involving sialic acid residues and suitably positioned hydroxyl groups in neighboring saccharide moieties is known to occur for glycolipids such as gangliosides. It has been suggested that these lactones are more immunogenic and tumor-specific than their native counterparts and that they might find use as cancer vaccines. We have now investigated if lactonization also occurs for the sialyl T_N and T antigens of mucins. It was found that the model compound sialyl T benzyl glycoside **6**, and the glycopeptide Ala-Pro-Asp-**Thr**-Arg-Pro-Ala **7** from the tandem repeat of the mucin MUC1, in which **Thr** stands for the 2,3-sialyl-T antigen, lactonized during treatment with glacial acetic acid. Compound **6** gave the 1″ → 2′ lactone as the major product and the corresponding 1″ → 4′ lactone as the minor product. For glycopeptide **7** the 1″ → 4′ lactone constituted the major product, whereas the 1″ → 2′ lactone was the minor one. When lactonized **7** was dissolved in water the 1″ → 4′ lactone underwent slow hydrolysis, whereas the 1″ → 2′ remained stable even after a 30 days incubation. In contrast the corresponding 2,6-sialyl-T_N glycopeptide **8** did not lactonize in glacial acetic acid.

Introduction

Mucins are members of an expanding family of large multifunctional glycoproteins present on the surface of many epithelial cells. Their main function is to provide lubrication and moisturisation of the surfaces of the epithelial tissues as well as protection against invasion of pathogenic microorganisms and mechanical injury.[1] Mucins are in general defined by the presence of a substantial amount of carbohydrates attached as *O*-glycans to threonines and serines and also by a high content of proline in their protein backbone. Of special interest is the membrane-bound mucin MUC1,[2,3] which is a heavily *O*-glycosylated, high-molecular weight glycoprotein present between many epithelia and their extracellular environments including those of the mammary gland, uterus, and gastrointestinal tract. Its protein backbone consists of repeating units of 20 amino acids with the sequence HGVTSAPDTRPAPGSTAPPA,[4] bearing five potential *O*-glycosylation sites. When MUC1 is produced by the mammary gland, galactosyl residues are attached to *N*-acetylgalactosamine on Ser/Thr leading to the core 1 structure Galβ(1 → 3)GalNAc α-Ser/Thr known as the T antigen. The T antigen serves as a substrate for the core 2 β1,6GlcNAc transferases forming the trisaccharide Galβ(1 → 3)(GlcNAcβ(1 → 6))GalNAc.[2,5] The latter can be further extended by adding *N*-acetyllactosamine to produce more complex and branched glycan side chains. Termination of the chain growth is accomplished by addition of sialic acid or fucose, or by sulfation.[6] In epithelial tumor cells, low expression of core 2 β1,6GlcNAc transferases combined with elevated levels of sialyl transferases give simplified carbohydrate patterns which may be accessible to the immune system. The T_N (GalNAcα-Ser/Thr), T (Galβ(1 → 3)GalNAcα-Ser/Thr), 2,6-sialyl-T_N (Neu5Acα(2 → 6)GalNAcα-Ser/Thr) and 2,3-sialyl-T (Neu5Acα(2 → 3)Galβ(1 → 3)GalNAcα-Ser/Thr) structures constitue tumor-associated antigens. The presence of these antigens on the surface of common malignant tumors has stimulated intense studies directed towards development of synthetic carbohydrate-based anticancer vaccines.[7–9]

For glycolipids, formation of intramolecular lactones is known to occur between sialic acid residues and suitably placed hydroxyl groups in neighboring galactose moieties.[10–12] It has been suggested that such lactones are more immunogenic and tumor-specific as compared to their native, open form and that they therefore have potential as immunogenes in development of cancer vaccines.[13] Since glycopeptides containing the 2,6-sialyl-T_N and 2,3-sialyl-T antigenes could potentially form lactones in acidic environment, there is a possibility that these structures could be immunological analogues to ganglioside lactones and constitue a starting point for development of glycopeptide based anticancer vaccines. In this article we describe synthesis and lactonization studies of the 2,6-sialyl-T_N and 2,3-sialyl-T antigens attached to a peptide from the MUC1 repeating unit. Methods for synthesis of tumor-associated glycopeptides from mucins have been reviewed recently.[14–17] In spite of the recent, successful chemical sialylation used in the syntheses of building blocks corresponding to the 2,6-sialyl-T_N[18,19] and 2,3-sialyl-T[20] antigens, low yields and poor stereoselectivity are often encountered due to the lack of neighboring group assistance on the sialyl donor. A convenient alternative is the chemoenzymatic approach to sialylated T and T_N antigens, using

[a]*Organic Chemistry, Department of Chemistry, Umeå University, SE-901 87, Umeå, Sweden. E-mail: mikael.elofsson@chem.umu.se, jan.kihlberg@chem.umu.se*
[b]*Department of Medicinal Biochemistry and Biophysics, Umeå University, SE-901 87, Umeå, Sweden*
[c]*Institute of Molecular Pathology & Immunology, University of Porto, 4200, Porto, Portugal*
[d]*AstraZeneca R & D Mölndal, SE-431 83, Mölndal, Sweden*

various enzymes which are readily available today. Depending on their specificity, these enzymes catalyze the transfer of a sialic acid moiety from cytidine-5′-monophospho-*N*-acetylneuraminic acid (CMP-Neu5Ac) in an *α*-selective fashion to a particular hydroxyl group of galactose residues. Consequently sialyl transferases were succesfully employed in the synthesis of wide range of oligosaccharides[21] and glycopeptides[22–24] operating with complete stereo- and regiocontrol. We therefore decided to use a chemoenzymatic approach for synthesis of the glycopeptides used in this article.

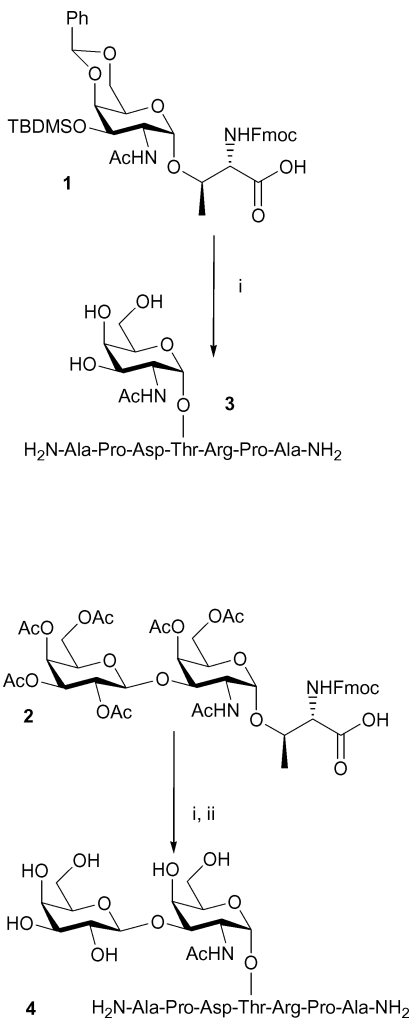
Results and discussion

Chemoenzymatic synthesis of glycopeptides

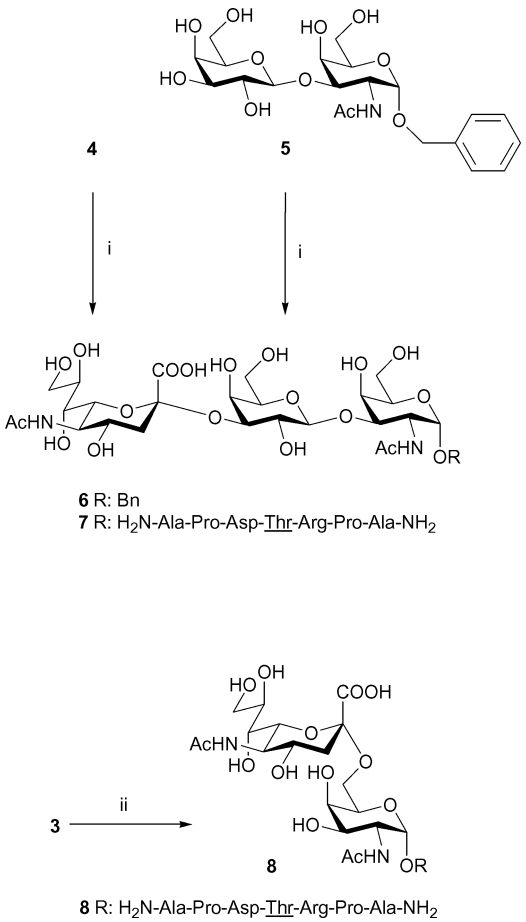
Synthesis of the T_N building block **1** (Scheme 1), carrying acid labile protective groups on the carbohydrate moiety, was accomplished from 4-methylphenyl 2-azido-2-deoxy-1-thio-β-D-galactopyranoside, according to a previously published procedure.[23] Silver triflate mediated glycosylation of Fmoc-3-*O*-(2-azido-4,6-*O*-benzylidene-2-deoxy-α-D-galactopyranosyl)-L-threonine *tert*-butyl ester with peracetylated galactosyl bromide, reductive acetylation of azide group, and deprotection of the *tert*-butyl ester afforded the T building block **2**[25] in 35% yield (Scheme 1). Both building blocks are suitable for use in solid-phase peptide synthesis.

After the synthesis of building blocks **1** and **2** we turned our attention to solid-phase synthesis of glycopeptide **3** and **4** (Scheme 1), that are based on the tandem repeating unit from the mucin MUC1 with the glycosylated threonine located in the center of the immunodominant region.[5] Incorporation of the two units **1** and **2** was performed using the standard Fmoc protocol for solid-phase peptide synthesis on Tentagel and ArgoGel resins. In the synthesis of **3** the acid labile protective groups originating from building block **1** were removed during the cleavage from the solid support, whereas glycopeptide **4** was deacetylated after cleavage from the resin using conditions previously described by us.[23,26]

With sufficient amounts of glycopeptides **3** and **4** in hand we turned our interest towards the extension of the carbohydrate side-chains by means of enzymatic sialylation. We thought that benzyl T glycoside **5**[27] could serve as a good model compound for sialylation using commercially available recombinant α2,3OST sialyl transferase from rat liver. Following standard sialylation procedures,[28,29] saccharide **5** was incubated with the sialic acid donor CMP-NeuAc and recombinant α2,3OST in buffer at pH 6 in preparative scale (Scheme 2). It was found that the enzyme



Scheme 1 Solid-phase synthesis of glycopeptides **3** and **4**. *Reagents and conditions*: (i) glycopeptide synthesis according to the Fmoc protocol. (ii) Deacetylation with NaOMe/MeOH, 20 mM.



Scheme 2 Enzymatic sialylation of benzyl glycoside **5** and glycopeptides **3** and **4**. *Reagents and conditions*: (i) CMP-Neu5Ac (1.75 equiv), α2,3OST, 0.25 M sodium cacodylate buffer (pH 6) containing 0.1% Triton X-100, 37 °C. (ii) CMP-Neu5Ac (2 mM), ST6GalNAc-I, 20 mM Bis-Tris buffer (pH 6.5) containing EDTA (20 mM) and dithiothreitol (1 mM), 37 °C.

transferred a sialic acid unit regio- and stereoselectively onto the 3-OH group of the terminal galactose moiety. After purification on a Waters Sep-Pak C18 solid-phase extraction column and reversed-phase HPLC the sialylated T benzyl glycoside **6** was produced in 72% isolated yield. In the sialylation more than 95% of the substrate **5** was consumed. The identity of compound **6** was confirmed by electrospray (ES) and fast atom bombardment (FAB) mass spectrometry and its ^{1}H NMR data were in agreement with those published.[30] This result promped us to investigate enzymatic sialylation of the T glycopeptide **4** and also of the corresponding T_N glycopeptide **3** (Scheme 2). The synthesis of glycopeptide **7** from **4** revealed that α-(2 → 3)-sialylation of glycopeptides which bear the T antigen is facile. Glycopeptide **4** was sialylated as described for **5** and the sialyl T glycopeptide **7** was isolated in 60% yield with approximately 90% conversion of the starting material as determined by reversed-phase HPLC. When glycopeptide **3** was sialylated on preparative scale using recombinant ST6GalNAc-I[31] sialyl transferase, **8** was obtained in 60% yield after purification. The identity of the sialylated glycopeptides **7** and **8** was confirmed by ES mass spectrometry and ^{1}H NMR spectroscopy (Tables 1 and 2). In conclusion this chemoenzymatic synthesis approach effectively produced sialylated derivatives suitable as lactonization substrates.

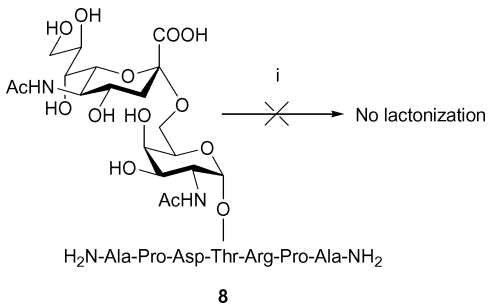
Formation of lactones

Lactonization studies of sialic acid residues of certain free oligosaccharides, *e.g.* 3′-sialyllactose[32] and several gangliosides including G_{M1},[12] G_{M3},[11] G_{M4},[10] G_{D1a},[33] and G_{D1b}[34] have been described. To the best of our knowledge lactonization of sialylated glycopeptides has not been investigated. Two methods have been applied in order to obtain intramolecular esters of different gangliosides; incubation in glacial acetic acid[11] or treatment with dicyclohexylcarbodiimide in anhydrous DMSO.[12] Under these conditions the ester linkage is often produced in almost quantitative yield. In this study we decided to use acetic acid incubation because of its ease of handling and the simplistic experimental procedure.

First, we turned our attention to the lactonization of free sialyl T antigen benzyl glycoside **6**. Compound **6** was dissolved in glacial deuterated acetic acid and allowed to incubate for 14 days at room temperature. The reaction was monitored every second day with ^{1}H NMR spectroscopy. Two lactones, **9a** and **9b** were produced in a 3 : 2 ratio (Scheme 3). In **9a** and **9b** the ester linkage was formed between the carboxyl group of *N*-acetylneuraminic acid and the C_2–OH and C_4–OH of galactose moiety, respectively. The conversion of **6** to **9a** and **9b** was complete according to the ^{1}H NMR spectra.

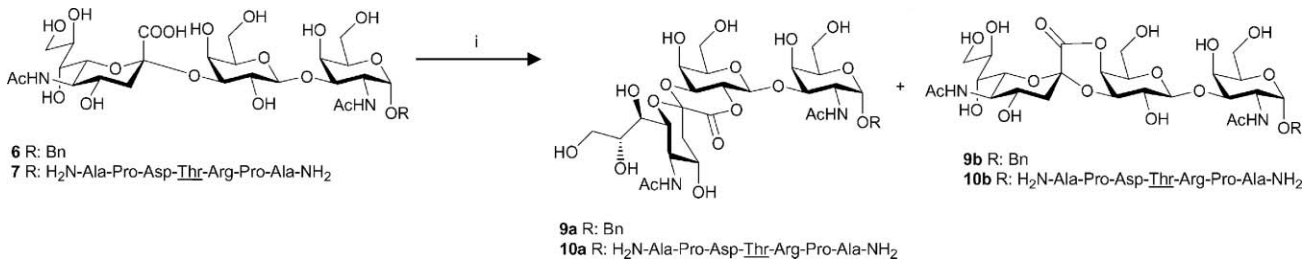
The most characteristic NMR features were significant high shifts (Table 1) for H-2 in galactose and H-4 of Neu5Ac in the lactone **9a**, and for H-4 in the galactose moiety of the lactone **9b** as compared with sialyl T glycoside **6**. Although all asignments were based on spectra in acetic acid-d_4, these type of large downfield shifts are comparable to those of lactone units acting as intermediates in the synthesis of sialyl Lewis^x derivatives,[35] the gangliosides G_{M1}, G_{M3}, and G_{M4},[10–12] and a glycosylated heptapeptide corresponding to the human M blood group determinant.[36]

Formation of intramolecular esters in sialyl T antigen glycoside **6** promped us to further investigate lactonization in sugar units of sialyl T and sialyl T_N glycopeptides **7** and **8**. First, sialyl T_N glycopeptide **8** was incubated in acetic acid-d_4 at ambient temperature for 14 days (Scheme 4). The reaction was monitored every second day using a set of standard 1D and 2D NMR experiments. However during the course of the reaction no changes in the NMR spectra were observed. This suggests that the formation of a seven-membered ring between the sialic acid carboxyl group and C_4–OH in galactosamine is not facile. Thereafter compound **7** was subjected to the lactonization conditions as decribed above during a time frame of 17 days (Scheme 3). The advance of the reaction was monitored each day using mainly 1D ^{1}H NMR experiments in combination with COSY, ROESY and TOCSY experiments. After a period of 11 days no further alterations in the ^{1}H NMR spectra were observed suggesting the culmination of the process. Although we were able to detect two different products by analysis of the NMR spectra recorded in acetic acid-d_4, it turned out to be difficult to assign resonances to a particular product. In order to make an accurate assignment the solvent was changed to DMSO-d_6. From a combination of hetero- and homonuclear experiments in DMSO-d_6, we could identify the two different products; the major product being the 1″ → 4′ lactone **10b** with the corresponding 1″ → 2′ lactone **10a** being formed as



8

Scheme 4 Lactonization of sialylated glycopeptide **8**. *Reagents and conditions*: (i) glacial acetic acid (AcOH-d_4), RT.



Scheme 3 Lactonization of sialylated benzyl glycoside **6** and sialylated glycopeptide **7**. *Reagents and conditions*: (i) glacial acetic acid (AcOH-d_4), RT.

Table 1 ^{1}H NMR chemical shifts (δ, ppm) for the the saccharide **6**[a], *O*-linked glycopeptides **7**[b] and **8**[c] and the lactones **9a**,[a] **9b**,[a] **10a**,[d] and **10b**[d]

	6	7	8	9a	9b	10a	10b
Benzyl							
H-a	4.61–4.64	—	—	n.d.[e]	n.d.[e]	—	—
H-b	4.85–4.88	—	—	n.d.[e]	n.d.[e]	—	—
Ar	7.39–7.52	—	—	7.41–7.55	7.41–7.55	—	—
GalNHAc							
N**H**Ac	n.d.[e]	7.81	7.57	n.d.[e]	n.d.[e]	n.d.[e]	n.d.[e]
H-1	5.14	3.98	3.94	5.17	5.17	n.d.[e]	n.d.[e]
H-2	4.64	4.19	4.05	4.68	4.68	n.d.[e]	n.d.[e]
H-3	4.17	3.73	3.82	4.17	4.17	n.d.[e]	n.d.[e]
H-4	4.43	n.d.[e]	3.57	4.47	4.47	n.d.[e]	n.d.[e]
Gal							
H-1	4.70	4.48	—	4.88	4.74	4.51	4.48
H-2	3.82	3.50	—	5.08	3.85	5.19	n.d [e]
H-3	4.26	4.01	—	4.12	4.32	4.10	4.06
H-4	4.14	3.71	—	4.17	5.51	n.d[e]	5.26
Neu5Ac							
N**H**Ac	n.d.[e]	8.04	8.01	n.d.[e]	n.d.[e]	n.d.[e]	n.d.[e]
H-3$_{ax}$	2.94	1.76	1.64	2.74	2.98	1.54	1.54
H-3$_{eq}$	2.07	2.72	2.66	1.96	2.12	2.30	2.27
H-4	4.12	3.65	3.65	4.52	4.13	4.13	4.14
H-5	3.83	3.82	3.80	4.22	n.d.[e]	3.50	3.53
H-6	n.d.[e]	n.d.[e]	n.d.[e]	n.d.[e]	n.d.[e]	n.d.[e]	3.24

[a] Recorded at 400 MHz, 298 K, with CD_3COOD ($\delta_H = 11.59$) as internal standard. [b] Recorded at 500 MHz, 298 K, with H_2O ($\delta_H = 4.75$) as internal standard containing 20% D_2O. [c] Recorded at 500 MHz, 298 K, with H_2O ($\delta_H = 4.75$) as internal standard containing 10% D_2O. [d] Recorded at 500 MHz, 298 K, with DMSO-d_6 ($\delta_H = 2.50$) as internal standard. [e] Not determined.

a minor product (ratio 7 : 3, Table 1). The relative ratio was based on the volumes of two crosspeaks derived from H-3 of neuraminic acid (Fig. 1). The key feature which allowed us to identify 1″ → 4′ lactone **10b** was the downfield peak at 5.26 ppm which was shown to be H-4 in the galactose moiety, in comparison with 3.71 ppm of H-4 of galactose in sialylated glycopeptide **7**. Downfield shifts of the same magnitude were also observed for lactones formed during synthesis of G_{M4}[10] and sialyl Lewis^X lactones.[37,38] The high shift peak at 5.19 ppm in lactone **10a** was assigned to H-2 in the galactose residue as compared to the peak at 3.50 ppm of H-2 in the non-lactonized glycopeptide **7**, thus showing extensive deshielding of the ester linkage. This large downfield shift of H-2 in galactose was also observed for 1″ → 2′ lactones in G_{M1}[12] and G_{M3}[11] gangliosides and for intermediate products in the synthesis of the sialyl Le^X trisaccharide.[39] As stated above, the assignment

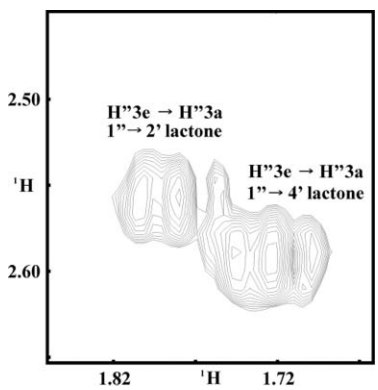


Fig. 1 Part of the TOCSY spectrum at the beginning of the hydrolysis. Two crosspeaks derived from H-3 equatorial (e) and axial (a) in the sialic acid residues of lactones **10a** and **10b**. (The spectrum was recorded in H_2O containing 10% D_2O).

of lactones **10a** and **10b** are derived from observed crosspeaks in COSY and TOCSY experiments. These lactonization effects are probably due to the reorientation of Neu5Ac carboxyl group after esterification and to the alteration in electronegativity along the Neu5Ac backbone. Although it was possible to identify two different peptide chains in post-lactonized mixture, we could not attribute them to a particular saccharide lactone.

Somewhat surprisingly the benzyl glycoside **6** gave approximately equal amounts of the 1″ → 2′ lactone **9a** and the 1″ → 4′ lactone **9b** while the 1″ → 4′ lactone **10b** dominated when glycopeptide **7** was lactonized. For lactones obtained from benzyl glycoside **6** this might be explained by the fact that no steric interactions occur between the sialic acid unit and the benzyl group, allowing the two possible lactones to be formed in relatively similar amounts. Construction of a space-filling model of the 1″ → 2′ lactone **10a** reveals that the Neu5Ac ring lies in a plane perpendicular to the plane of galactose–galactosamine moiety, hence interacting with the peptide chain. This could be the explanation for the fact that **10a** was produced as the minor product. On the other hand a model of the major 1″ → 4′ lactone **10b** shows that Neu5Ac backbone lies in the same plain as galactose and galactosamine, producing almost no contact with the peptide moiety.

Hydrolysis of lactone products 10a and 10b

It has been suggested that lactones originating from gangliosides are more immunogenic than the gangliosides themselves, implying that the lactones might be valuable as immunogens for vaccination against tumors.[13] However sialic acid lactones might not be sufficiently stable to be therapeutically useful. Having now in our hands the mixture of the two glycopeptide lactones **10a** and **10b**, we were therefore interested in investigating their hydrolytic

Table 2 ^{1}H NMR chemical shifts (δ, ppm) for the amino acids in glycopeptides **7**[a] and **8**[b]

	7	**8**
Ala[1]		
NH	8.030	n.d.[d]
α-H	4.329	4.328
β-H	1.521	1.512
Pro[2]		
α-H	4.460	4.461
β-H	1.866, 2.308	1.874, 2.322
γ-H	2.007[c]	2.013[c]
δ-H	3.604, 3.690	3.596, 3.698
Asp[3]		
NH	8.628	8.635
α-H	4.813	4.822
β-H	2.794, 2.914	2.812, 2.920
Thr[4]		
NH	8.640	8.378
α-H	4.504	4.496
β-H	4.366	4.270
γ-H	1.230	1.259
Arg[5]		
NH	8.334	8.389
α-H	4.534	4.552
β-H	1.688, 1.820	1.668, 1.821
γ-H	1.668[c]	1.668[c]
δ-H	3.193[c]	3.194[c]
δ-NH	7.232	7.170
Pro[6]		
α-H	4.332	4.314
β-H	1.897, 2.293	1.898, 2.296
γ-H	2.031[c]	2.020[c]
δ-H	3.595, 3.725	3.596, 3.725
Ala[7]		
NH	8.456	8.437
α-H	4.186	4.185
β-H	1.377	1.380

[a] Recorded at 500 MHz, 298 K, with H_2O ($\delta_H = 4.75$) as internal standard containing 20% D_2O. [b] Recorded at 500 MHz, 298 K, with H_2O ($\delta_H = 4.75$) as internal standard containing 10% D_2O. [c] Degeneracy has been assumed. [d] Not determined.

stability. The mixture of **10a** and **10b** was dissolved in H_2O/D_2O and the solution was kept at ambient temperature for 30 days with NMR experiments being run every second day. We found that the 1″ → 4′ lactone **10b** was completely hydrolysed to give the sialyl T peptide **7** after 9 days (Fig. 2). Interestingly, the 1″ → 2′ lactone **10a** showed extensive stability compared to the 1″ → 4′ lactone **10b** as it remained in the aqueous solution even after a

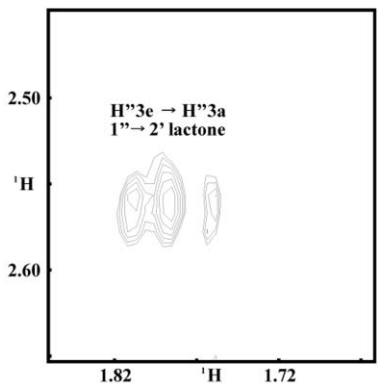


Fig. 2 Crosspeak derived from the residual lactone **10a** after 9 days in water. (The spectrum was recorded in H_2O containing 10% D_2O).

period of 30 days, as determined by ^{1}H NMR spectroscopy. This remarkable stability might be due to the highly rigid structure of **10a** where the Neu5Ac ring lies in a plane perpendicular to the plane of the galactose moiety and lactone ring. Since lactone **10a** showed reasonable stability in an aqueous environment, we suggest that it could serve as a good immunogen also under physiological conditions.

Lactams, which are hydrolytically more stable and structurally similar to lactones, should be good substitutes. The synthesis of ganglioside lactams corresponding to G_{M2}, G_{M3} and G_{M4} ganglioside lactones has been reported previously.[40–42] They proved to be very stable upon storage with only minor hydrolysis occuring in D_2O at 37 °C during one month. It was also found that antibodies raised towards G_{M3}-lactam cross-reacted with the G_{M3}-lactone *in vitro*.[43]

In conclusion, a chemoenzymatic approach has been used to efficiently prepare sialylated T and T_N glycopeptides derived from the mucin MUC1. By using these glycopeptides the formation of intramolecular lactones in sialylated glycopeptides was examined for the first time. The sialyl T glycopeptide formed 1″ → 4′ and 1″ → 2′ lactones upon treatment with acetic acid. The 1″ → 2′ lactone, in contrast to the 1″ → 4′ lactone, showed significant stability in aqueous solution, implying its potential for use as immunogen in cancer vaccines. No lactones were formed from the sialyl T_N glycopeptide, even after prolonged treatment with acetic acid.

Experimental

General methods and materials

All reactions were carried out under an inert nitrogen atmosphere using dry, freshly distilled solvents under anhydrous conditions, unless otherwise stated. CH_2Cl_2 was distilled from calcium hydride. Organic solutions were dried over Na_2SO_4 before being concentrated. TLC was performed on Silica Gel F_{254} (Merck) with detection by UV light and by charring with 10% sulfuric acid. Flash column chromatography was performed on Silica Gel (Matrex, 60 Å, 35–70 μm, Grace Amicon). Preparative reversed-phase HPLC was performed on a Kromasil C-8 column (250 × 20 mm, 5 μm, 100 Å), eluted with a linear gradient of MeCN in H_2O containing 0.1% TFA, with a flowrate of 11 mL min^{-1} and detection at 214 nm. Analytical HPLC was performed on a Beckman System Gold HPLC, using a Kromasil C-8 column (250 × 4.6 mm, 5 μm, 100 Å), with the same eluent flowrate of 1.5 mL min^{-1} and detection at 214 nm.

^{1}H and ^{13}C NMR spectra were recorded on a Bruker DRX-400 spectrometer and a Bruker AMX2-500 spectrometer (Massachusetts, USA). All NMR experiments were conducted at 298 K using $CDCl_3$ [residual $CHCl_3$ at 7.26 ppm (δ_H)], DMSO-d_6 [residual DMSO-d_5 at 2.50 ppm (δ_H) and 39.60 ppm (δ_C)], CD_3CO_2D [residual CD_3CO_2DH at 11.59 ppm (δ_H) and 20.0 ppm (δ_C)] or CD_3OD [residual CD_2HOD at 3.35 ppm (δ_H)]. The chemical shift of the water signal was used as a reference for compounds dissolved in water and calibrated to 4.75 ppm. The ^{13}C HSQC spectra were calibrated using the gyromagnetic ratio for carbon.[44] The spectra used for resonance assignments included phase sensitive DQF-COSY,[45] TOCSY,[46] ROESY[47] and gradient enhanced HSQC.[48] The DIPSI pulse sequence

with a spin lock time of 45 or 75 ms was used in the TOCSY experiments whereas the ROESY spectra were recorded with a mixing time of 200 ms. Mass spectra were recorded on a Waters Micromass ZQ using positive electrospray ionization (ES+). High-resolution fast atom bombardment mass spectra (HRMS) were recorded with a JEOL SX102 A mass spectrometer. Ions for FABMS were produced by a beam of xenon atoms (6 keV) from a matrix of glycerol and thioglycerol. (2-acetamido-2-deoxy-3-*O*-*β*-D-galactopyranosyl-*α*-D-galactopyranosyl)-*O*-benzyl glycoside **5** (benzyl T antigen) was purchased from Calbiochem® (Germany). The T_N antigen N^{α}-fluoren-9-ylmethoxycarbonyl-3-*O*-(2-acetamido-4,6-*O*-benzylidene-3-*O*-*tert*-butyldimethylsilyl-2-deoxy-α-D-galactopyranosyl)-L-threonine **1** was prepared as described previously.[23] The threonine-based T antigen, N^{α}-fluoren-9-ylmethoxycarbonyl-*O*-[*O*-(2′,3′,4′,6′-tetra-*O*-acetyl-*β*-D-galactopyranosyl)-(1′ → 3)-2-acetamido-4,6-di-*O*-acetyl-2-deoxy-*α*-D-galactopyranosyl]-L-threonine was synthesised accordning to previously reported procedure.[25] Glycosylated building blocks had data in agreement with the cited references.

General procedure for solid-phase peptide synthesis

Glycopeptides were prepared under conditions identical to those described previously.[49] A Tentagel S NH_2 resin (Rapp Polymere, Germany) functionalized with the Rink amide linker {*p*-[α-fluoren-9-ylmethoxyformamido)-2,4-dimethoxybenzyl]-phenoxyacetic acid}[50,51] B(achem AG, Switzerland) and ArgoGel®-Rink-NH-Fmoc resin (Argonaut Technologies Inc.) were used in the synthesis of glycopeptides. N^{α}-Fmoc-amino acids (Neosystem, France and Bachem, Switzerland) with the following side chain protecting groups were used in the synthesis: 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Pmc) for arginine and *tert*-butyl (*t*Bu) for aspartic acid. DMF was distilled before being used. Couplings were performed manually in a mechanically agitated reactor. Fmoc amino acids (4 equiv) were activated as benzotriazolyl esters by using 1,3-diisopropylcarbodiimide (DIC, 3.9 equiv) and 1-hydroxybenzotriazole (HOBt, 6 equiv) in dry DMF. Acylations were monitored by using bromophenol blue (BFB, 2 mM solution in DMF) as indicator[52] (the colour of the reaction mixture changes from blue to yellow). N^{α}-Fmoc deprotections were effected by a flow of 20% solution of piperidine in DMF for 3 min and futher by shaking for 7 min. Before and after treatment with piperidine solution the resin was washed five times with DMF. The threonine-based building blocks: **1** (1 equiv, 40 µmol) and **2** (1.04 equiv, 52.2 µmol) were activated in distilled DMF (1 mL) at room temperature during 1–2 min by addition of 1,3-diisopropylcarbodiimide (1.1 equiv) and 1-hydroxy-7-azabenzotriazole (HOAt, 3.3 equiv). The activated esters were then coupled to the peptides resin during 24 h.

After completion of the synthesis, the resin carrying the protected glycopeptide was washed with CH_2Cl_2 (five times) and dried under vacuum. Cleavage from the resin and removal of acid labile protective groups was performed with TFA/H_2O/thioanisole/ethanedithiol (87.5 : 5 : 5 : 2.5, 20 mL/200 mg of resin) for 3.5 h, followed by filtration. Acetic acid (10 mL) was added to filtrate which was then concentrated, and acetic acid was added again (3 × 10 mL) followed by concentration after each addition. The residue was triturated with cold ethyl ether (three times), dissolved in a mixture of water and acetic acid (6 : 1), and then freeze-dried.

Deprotection of acetates was performed with methanolic sodium methoxide solution (0.02 M, 3 mL/8 mg in 8 mL MeOH), under nitrogen atmosphere for several hours, with careful monitoring on HPLC. Solutions were neutralized with acetic acid and concentrated under vacuum. The residues were then purified by reversed-phase HPLC.

L-Alanyl-L-prolyl-L-aspartyl-3-*O*-(2-acetamido-2-deoxy-3-*O*-*β*-D-galactopyranosyl-*α*-D-galactopyranosyl)-L-threonyl-L-arginyl-L-prolyl-L-alanine amide (4). Synthesis, cleavage of the resin-bound glycopeptide, and then purification by reversed-phase HPLC (gradient 0 → 80% CH_3CN in H_2O, both containing 0.1% TFA, during 60 min) gave the acetylated target glycopeptide [22 mg, MS (ES): calcd 1343.3, found 1343.4]. The acyl protected glycopeptide (8 mg) was subjected to deacetylation with methanolic sodium methoxide which was carried out for 6 h (careful monitoring with HPLC). Purification by preparative HPLC (gradient 0 → 30% CH_3CN in H_2O, with 0.1% TFA, during 60 min) gave **4** (4.6 mg, 25% overall yield). MS (ES): calcd for $C_{44}H_{74}N_{12}O_{20}$ 1091.522 *m*/*z* $(M + H)^+$, observed 1091.647.

L-Alanyl-L-prolyl-L-aspartyl-3-*O*-(2-acetamido-2-deoxy-*α*-D-galactopyranosyl)-L-threonyl-L-arginyl-L-prolyl-L-alanine amide (3). Synthesis, cleavage of the resin-bound glycopeptide together with deprotection of acid labile protective groups on sugar moiety, and then purification by reversed-phase HPLC (isocratic elution 5% CH_3CN in H_2O, both containing 0.1% TFA, during 30 min) gave the target glycopeptide **3** (8.8 mg, 24% yield). MS (FAB): calculated for $C_{38}H_{64}N_{12}O_{15}$ 929.468 *m*/*z* $(M + H)^+$, observed 929.447.

General procedure for enzymatic sialylation of the compounds 4 and 5

Soluble recombinant form of *α*2,3OST *Spodoptera frugiperda* (76–81 m units, *β*-D-galactosyl-*β*-1,3-*N*-acetyl-*β*-D-galactosamine-*α*-2,3-sialyltransferase, Calbiochem®, Germany) from rat liver (≥1 unit mg^{-1} protein) was used for sialylation of glycopeptide **4** and benzyl T glycoside **5**. The reactions were performed in 0.25 M sodium cacodylate buffer (2 mL, pH 6) containing 0.1% Triton X-100 and CMP-*β*-D-sialic acid disodium salt donor (1.75 equiv, Calbiochem®, Germany). The solution was then incubated at 37 °C for 1–2 days. Progress of the reaction was monitored by TLC. Purification on pre-conditioned Waters Sep-Pak Vac 12cc C-18 solid-phase extraction column (2 g) and then by preparative reversed-phase HPLC gave sialylated compounds.

Benzyl [(5-acetamido-3,5-dideoxy-D-glycero-*α*-D-galacto-2-nonulopyranosylonic acid)-(2 → 3)-*O*-*β*-D-galactopyranosyl]-(1 → 3)-2-acetamido-2-deoxy-*α*-D-galactopyranoside (6). The reaction progress was monitored with TLC (EtOH : AcOH : MeOH : H_2O 6 : 3 : 3 : 2). Purification on Waters Sep-Pak C18 solid-phase extraction column with gradient elution 0 → 50% MeOH in H_2O and futher by preparative HPLC (gradient 0 → 30% CH_3CN in H_2O during 30 min) gave **6** (2.4 mg, 72% yield). MS (FAB): calcd for $C_{32}H_{48}N_2O_{19}$ *m*/*z* $(M + 2Na–H)^+$ 809.2568, observed 809.2565. ^1H NMR data was in agreement with that published previously.[30]

L-Alanyl-L-prolyl-L-aspartyl-3-*O*-[(5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid)-(2 → 3)-*O*-β-D-galactopyranosyl-(1 → 3)-*O*-(2-acetamido-2-deoxy-α-D-galacto-pyranosyl)]-L-threonyl-L-arginyl-L-prolyl-L-alanine amide (7). Synthesis, purification on Waters Sep-Pak C18 solid-phase extraction column with gradient elution 0 → 50% MeOH in H_2O and futher by preparative HPLC (gradient 0 → 10% CH_3CN in H_2O during 30 min, then 10 → 100% during 30 min) gave **7** (3 mg, 60% yield). MS (ES): calcd for $C_{55}H_{91}N_{13}O_{28}$ m/z $(M + H)^+$ 1382.618, observed 1382.617.

L-Alanyl-L-prolyl-L-aspartyl-3-*O*-[(5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid)-(2 → 6)-*O*-(2-acetamido-2-deoxy-α-D-galactopyranosyl)]-L-threonyl-L-arginyl-L-prolyl-L-alanine amide (8). Sialylation of **3** was essentially performed as previously described:[23] compound **3** (3 mg) was added to purified ST6GalNAc-I (12 m units)[53] in 20 mM Bis-Tris buffer (pH 6.5) containing CMP-β-D-sialic acid disodium salt donor (2 mM), EDTA (20 mM), and dithiothreitol (1 mM). The solution was then incubated at 37 °C for 6 h, and monitoring of sialylation was performed by nanoscale reverse-phase HPLC[54] in combination with MALDI-TOF mass spectrometry. Purification by preparative HPLC (gradient 0 → 80% CH_3CN in H_2O in 60 min) gave **8** (2.37 mg, 60% yield). MS (ES): calcd for $C_{49}H_{81}N_{13}O_{23}$ m/z $(M)^+$ 1220.24, observed 1220.25.

General procedure for the lactonization

The sialylated benzyl glycoside **6** and glycopeptides **7** and **8** were dissolved in CD_3COOD and incubated at ambient temperature. The reactions were monitored using mainly ^1H NMR spectroscopy over a period of 14 days in order to verify completion. The product mixture **10a** and **10b** was freeze-dried several times from acetic acid, dissolved in DMSO-d_6 and the set of NMR experiments were recorded, which allowed us to elucidate their structures.

Hydrolysis of mixture of lactones 10a and 10b

The lactones **10a** and **10b** were freeze-dried several times from DMSO and dissolved in D_2O. Hydrolysis was performed in NMR tube at ambient temperature and several NMR experiments were run every second day during 30 days in order to verify the reaction progress. The NMR spectra in DMSO-d_6 were also recorded after multiple lyophilisation from water.

Acknowledgements

This work was funded by the EU Prime Boost grant (QLRT-2001-02010). Celso A. Reis was supported by FCT (POCTI/CBO/44598/02) and the Association for International Cancer Research (Grant 05–088).

References

1 Y. S. Kim, J. Gum and I. Brockhausen, *Glycoconjugate J.*, 1996, **13**, 693–707.
2 F.-G. Hanisch and S. Müller, *Glycobiology*, 2000, **10**, 439–449.
3 J. Taylor-Papadimitriou, J. M. Burchell, T. Plunkett, R. Graham, I. Correa, D. Miles and M. Smith, *J. Mammary Gland Biol. Neoplasia*, 2002, **7**, 209–221.
4 S. J. Gendler, C. A. Lancaster, J. T. Papadimitriou, T. Duhig, N. Peat, J. Burchell, L. Pemberton, E. N. Lalani and D. Wilson, *J. Biol. Chem.*, 1990, **265**, 15285–15293.
5 J. Taylor-Papadimitriou, J. Burchell, D. W. Miles and M. Dalziel, *Biochim. Biophys. Acta*, 1999, **1455**, 301–313.
6 I. Brockhausen, *Biochem. Soc. Trans.*, 2003, **31**, 318–325.
7 S. J. Danishefsky and J. R. Allen, *Angew. Chem., Int. Ed.*, 2000, **39**, 836–863.
8 T. Toyokuni and A. K. Singhal, *Chem. Soc. Rev.*, 1995, 231–242.
9 G. Ragupathi, D. M. Coltart, L. J. Williams, F. Koide, E. Kagan, J. Allen, C. Harris, P. W. Glunz, P. O. Livingston and S. J. Danishefsky, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 13699–13704.
10 T. Terabayashi, T. Ogawa and Y. Kawanishi, *J. Biochem.*, 1990, **107**, 868–871.
11 R. K. Yu, T. A. W. Koerner, S. Ando, H. C. Yohe and J. H. Prestegard, *J. Biochem.*, 1985, **98**, 1367–1373.
12 G. Kirschner, G. Fronza, H. Egge, R. Ghidoni, D. Acquotti, G. Tettamanti and S. Sonnino, *Glycoconjugate J.*, 1985, **2**, 343–354.
13 G. A. Nores, T. Dohi, M. Taniguchi and S.-I. Hakomori, *J. Immunol.*, 1987, **139**, 3171–3176.
14 M. Meldal, in *Neoglycoconjugates: Preparation and Applications*, ed. Y. C. Lee and R. T. Lee, Academic Press, Inc., San Diego, 1994, pp. 145–198.
15 T. Norberg, B. Lüning and J. Tejbrant, *Methods Enzymol.*, 1994, **247**, 87–106.
16 C. Brocke and H. Kunz, *Bioorg. Med. Chem.*, 2002, **10**, 3085–3112.
17 G.-J. Boons and A. V. Demchenko, *Chem. Rev.*, 2000, **100**, 4539–4565.
18 C. Brocke and H. Kunz, *Synthesis*, 2004, 525–542.
19 M. Elofsson, L. A. Salvador and J. Kihlberg, *Tetrahedron*, 1997, **53**, 369–390.
20 S. Dziadek, C. Brocke and H. Kunz, *Chem. Eur. J.*, 2004, **10**, 4150–4162.
21 O. Blixt, K. Allin, L. Pereira, A. Datta and J. C. Paulson, *J. Am. Chem. Soc.*, 2002, **124**, 5739–5746.
22 C. Unverzagt, S. Kelm and J. C. Paulson, *Carbohydr. Res.*, 1994, **251**, 285–301.
23 S. K. George, T. Schwientek, B. Holm, C. A. Reis, H. Clausen and J. Kihlberg, *J. Am. Chem. Soc.*, 2001, **123**, 11117–11125.
24 Y. Takano, N. Kojima, Y. Nakahara, H. Hojo and Y. Nakahara, *Tetrahedron*, 2003, **59**, 8415–8427.
25 N. Mathieux, H. Paulsen, M. Meldal and K. Bock, *J. Chem. Soc., Perkin Trans. 1*, 1997, 2359–2368.
26 P. Sjölin, M. Elofsson and J. Kihlberg, *J. Org. Chem.*, 1996, **61**, 560–565.
27 Y. C. Lee, N. Kojima, E. Wada, N. Kurosawa, T. Nakaoka, T. Hamamoto and S. Tsuji, *J. Biol. Chem.*, 1994, **269**, 10028–10033.
28 B. Ernst and R. Oehrlein, *Glycoconjugate J.*, 1999, **16**, 161–170.
29 O. Schwardt, G.-P. Gao, T. Visekruna, S. Rabbani, E. Gassmann and B. Ernst, *J. Carbohydr. Chem.*, 2004, **23**, 1–26.
30 J. Xia, J. L. Alderfer, C. F. Piskorz, R. D. Locke and K. L. Matta, *Carbohydr. Res.*, 2000, **328**, 147–163.
31 N. Kurosawa, S. Takashima, M. Kono, Y. Ikehara, M. Inoue, Y. Tachida, H. Narimatsu and S. Tsuji, *J. Biochem. (Tokyo)*, 2000, **127**, 845–854.
32 W. A. Bubb, T. Saito, I. Arai, T. Urashima and T. Nakamura, *Carbohydr. Res.*, 2000, **329**, 471–476.
33 G. Fronza, G. Kirschner, D. Acquotti, R. Bassi, L. Tagliavacca and S. Sonnino, *Carbohydr. Res.*, 1988, **182**, 31–40.
34 D. Acquotti, G. Fronza, L. Riboni, S. Sonnino and G. Tettamanti, *Glycoconjugate J.*, 1987, **4**, 119–127.
35 K. C. Nicolaou, C. W. Hummel, N. J. Bockovich and C.-H. Wong, *J. Chem. Soc., Chem. Commun.*, 1991, 870–872.
36 Y. Nakahara, H. Iijima and T. Ogawa, *Tetrahedron Lett.*, 1994, **35**, 3321–3324.
37 U. Sprengard, G. Kretzschmar, E. Bartnik, C. Hüls and H. Kunz, *Angew. Chem., Int. Ed. Engl.*, 1995, **34**, 990–993.
38 W. Stahl, U. Sprengard, G. Kretzschmar and H. Kunz, *Angew. Chem., Int. Ed. Engl.*, 1994, **33**, 2096–2098.
39 S. J. Danishefsky, J. Gervay, J. M. Peterson, F. E. McDonald, K. Koseki, D. A. Griffith, T. Oriyama and S. P. Marsden, *J. Am. Chem. Soc.*, 1995, **117**, 1940–1953.
40 A. K. Ray, A. Nilsson and G. Magnusson, *J. Am. Chem. Soc.*, 1992, **114**, 2256–2257.
41 M. Wilstermann, L. O. Kononov, U. Nilsson, A. K. Ray and G. Magnusson, *J. Am. Chem. Soc.*, 1995, **117**, 4742–4754.

42 M. Wilstermann and G. Magnusson, *J. Org. Chem.*, 1997, **62**, 7961–7971.
43 K. Ding, A. Rosen, A. K. Ray and G. Magnusson, *Glycoconjugate J.*, 1992, **9**, 303–306.
44 D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley and B. D. Sykes, *J. Biomol. NMR*, 1995, **6**, 135–140.
45 M. Rance, O. W. Sorensen, G. Bodenhausen, G. Wagner, R. R. Ernst and K. Wüthrich, *Biochem. Biophys. Res. Commun.*, 1983, **117**, 479–485.
46 L. Braunschweiler and R. R. Ernst, *J. Magn. Reson.*, 1983, **53**, 521–528.
47 A. A. Bothner-By, R. L. Stephens, J.-M. Lee, C. D. Warren and R. W. Jeanloz, *J. Am. Chem. Soc.*, 1984, **106**, 811–813.
48 L. E. Kay, P. Keifer and T. Saarinen, *J. Am. Chem. Soc.*, 1992, **114**, 10663–10665.
49 J. Broddefalk, M. Forsgren, I. Sethson and J. Kihlberg, *J. Org. Chem.*, 1999, **64**, 8948–8953.
50 H. Rink, *Tetrahedron Lett.*, 1987, **28**, 3787–3790.
51 M. S. Bernatowicz, S. B. Daniels and H. Köster, *Tetrahedron Lett.*, 1989, **30**, 4645–4648.
52 M. Flegel and R. C. Sheppard, *J. Chem. Soc., Chem. Commun.*, 1990, 536–538.
53 N. T. Marcos, S. Pinho, C. Grandela, A. Cruz, B. Samyn-Petit, A. Harduin-Lepers, R. Almeida, F. Silva, V. Morais, J. Costa, J. Kihlberg, H. Clausen and C. A. Reis, *Cancer Res.*, 2004, **64**, 7050–7057.
54 J. Gobom, E. Nordhoff, E. Mirgorodskaya, R. Ekman and P. Roepstorff, *J. Mass Spectrom.*, 1999, **34**, 105–116.

