Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 339 (2004) 2329–2335
Carbohydrate RESEARCH

Note

Human serum IgA1 is substituted with up to six O-glycans as shown by matrix assisted laser desorption ionisation time-of-flight mass spectrometry

Edward Tarelli,[a,*] Alice C. Smith,[b] Bruce M. Hendry,[c] Stephen J. Challacombe[c] and Shideh Pouria[c]

[a]*Medical Biomics Centre, St. George's Hospital Medical School, Cranmer Terrace, London SW17 ORE, UK*
[b]*Department of Infection, Immunity and Inflammation. University of Leicester, LE5 4PW, UK*
[c]*GKT School of Medicine and Dentistry. London SE5 9RS, UK*

Received 17 May 2004; received in revised form 16 July 2004; accepted 18 July 2004
Available online 14 August 2004

Abstract—The micro-heterogeneity of human serum IgA1 results from variable O-glycan substitutions in the 'hinge region' of the molecule and this O-glycosylation may be altered in a number of medical conditions. This micro-heterogeneity has been monitored by analysis of IgA1-derived tryptic O-glycopeptides using matrix assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI-ToF-MS) analysis. With ammonium citrate–trihydroxyacetophenone matrix, individual compositional glycoforms have been baseline resolved in more than 70 samples and these spectra revealed for the first time that, in addition to expected substitution with 3,4 and 5 GalNAcs, a sixth GalNAc substitution was also present in the hinge region of the molecule. The spectra obtained from subsequent exoglycosidase-treated samples confirmed hexa-O-substitution. Following endoprotease digestions of the exoglycosidase treated samples, possible locations for the sixth GalNAc were indicated from further MALDI-ToF-MS analysis. Hexa-substitution accounts for around 5–10% the glycoforms. This is, we believe, the first report of hexa-O-substitution with GalNAc of human serum IgA1.

Keywords: MALDI-ToF-MS; Glycoprotein; Exoglycosidase; Hinge glycopeptide; Immunoglobulin A1; O-glycosylation; Endoprotease

Human serum IgA1 possesses two N-glycosylation sites,[1,2] and unusually for a circulating glycoprotein, an O-glycosylated proline-rich hinge region[3,4] (Fig. 1). In the IgA1 hinge region of both myeloma and normal serum, up to five O-glycosylation sites have been identified by amino acid sequencing and mass spectrometry.[1,4,5] For normal serum IgA1, Thr 228, Ser 230 and Ser 232 are reported to be fully O-glycosylated, while Thr 225 and Thr 236 are reported to be partially substituted, by O-GalNAc residues that in turn may be substituted by further galactose and neuraminic acid residues (Fig. 1).[4] It has been suggested that altered O-glycosylation may have a role in the deposition of IgA immune complexes in renal diseases such as IgA nephropathy.[6] Insights into the structure of IgA1 should therefore allow meaningful comparisons with diseased states and in the elucidation of their pathogenesis. The O-glycan composition of IgA1 has been previously profiled with some success by MALDI-ToF-MS of the derived O-glycopeptides using either α-cyano-4-hydroxy-cinnamic acid or dihydroxybenzoic acid/5-methoxysalicylic acid matrix.[5,7,8] These studies have revealed species in which up to five of the nine potential sites available for O-glycosylation are substituted. Recently we reported an

Abbreviations: MALDI-ToF-MS, matrix assisted laser desorption ionisation time-of-flight mass spectrometry; HGP, hinge glycopeptide; THAP, 2,4,6-trihydroxyacetophenone.

* Corresponding author. Tel.: +44-020-8725-5144; fax: +44-020-8725-0245; e-mail: etarelli@sghms.ac.uk

0008-6215/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2004.07.011

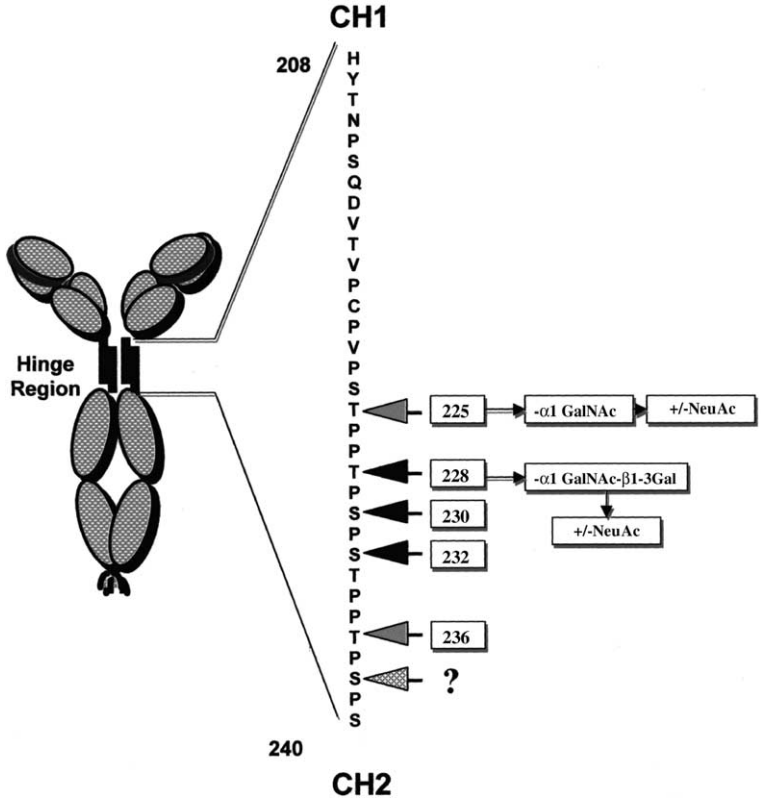

Figure 1. Schematic of the serum IgA1 molecule indicating the hinge region and the 33mer (208–240) hinge glycopeptide isolated after trypsin digestion. The black arrows demarcate the fully occupied positions, and the grey arrows the partially occupied positions of O-glycosylation in normal serum IgA1. The sixth glycan (?) may be located at any of the remaining S or T residues. Possible O-glycan permutations that may be present at any of these sites are as indicated.

improved methodology,[9] which uses 2,4,6-trihydroxyacetophenone–ammonium citrate (THAP) matrix.[10] THAP matrix provides superior baseline resolution of the different compositional O-glycopeptides present in a pool, enabling the degree of O-substitution to be readily determined from the observed *m*/*z* values and for detailed analyses to be performed including identification of low abundance species.

The tryptic fragment isolated from IgA1 that was used in the study presented here is the 33mer (208) HYTNPSQDVTVPCPVPSTPPTPSPSTPPTPSPS(240) (residue numbers are those of intact IgA1) and is a result of an unusual trypsin cleavage on the C-terminal side of serine 240, which has been attributed[7] to the presence, in the reduced and alkylated protein, of two adjacent pyridylethylated cysteines (residues 241 and 242). The O-glycopeptides (hinge glycopeptides, HGP) were prepared for analysis by MALDI-ToF-MS, which was performed both on the intact HGP pools and also on seven of the pools, modified by exoglycosidase digestions. These were treated sequentially with neuraminidase (EC 3.2.1.18), β-1→3 galactosidase (EC 3.2.1.23) and α-*N*-acetylgalactosaminidase (EC 3.2.1.49). Figure 2a shows a typical MALDI-ToF mass spectrum, obtained from a pool of intact IgA1 hinge glycopeptide. Similar multi-component spectra have been observed from all other HGP samples analysed in the course of our studies (data not shown). More than 30 peaks are discernible in Figure 2a, and their *m*/*z* values can be assigned to specific glycan content, providing a compositional profile of the glycoforms present (Table 1). For example, the species observed at *m*/*z* 6015 can be assigned to glycopeptide 543, that is the 33mer peptide substituted with 5 GalNAc, 4 Gal and 3 NeuAc residues (calculated *m*/*z* 6014; $[M+H]^+$ average masses are used throughout the text) whilst that at *m*/*z* 5358 can be assigned to glycopeptide 432 (calculated *m*/*z* 5358). Most peaks above background could be attributed to tri-, tetra- and penta-O-glycosylated species (Table 1). The spectrum also contained a number of minor peaks that could not be assigned in this manner, but their *m*/*z* values did however correspond to species that were hexa-O-glycosylated. These were (Fig. 2a and Table 1) observed at *m*/*z* 6672, 6510, 6382, 6218, 6056, 5928, 5765 and corresponded to species 654 (calculated 6671), 644 (calculated 6509), 653 (calculated 6380), 643 (calculated 6217), 633 (calculated 6055), 642 (calculated 5926) and 632 (calculated 5764). Hexa-O-substitution was corroborated

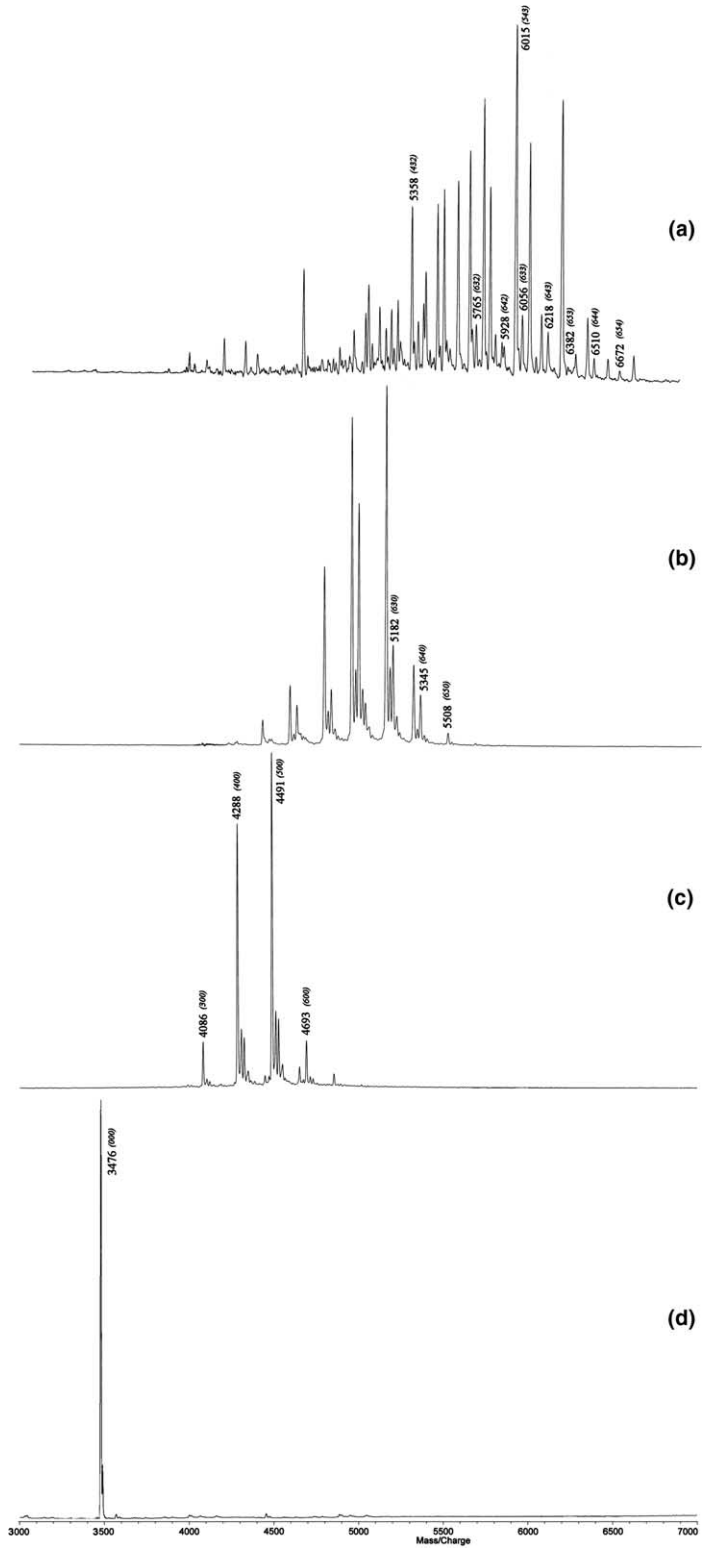

Figure 2. MALDI-ToF-MS of (a) the pool of intact hinge glycopeptides prepared from serum IgA1; (b) the pool from (a), following treatment with neuraminidase; (c) the pool from (b), following treatment with β-1→3 galactosidase; (d) the pool from (c), following treatment with α-*N*-acetylgalactosaminidase. *m*/*z* values of selected peaks are shown together, in parenthesis and italicised, with a three number code indicating the glycan composition of that glycopeptide, for example, 654 indicating 6 GalNAc, 5 Gal and 4 NeuAc. In (a) and (b), only 6 × GalNAc-substituted (with in (a), one example each of 5 × GalNAc-substituted and 4 × GalNAc-substituted) glycopeptides are annotated for reasons of clarity. Relative peak intensities (% of total) are shown in Table 1.

Table 1. Mass list of peaks observed in Figure 2a,b and c, their assignment and their % of total peak area determined by the Java program[9]

Figure 2a			Figure 2b			Figure 2c		
Mass	Glycan	%Total	Mass	Glycan	%Total	Mass	Glycan	%Total
4084	300	0.3	4411	320	1.2	4086	300	2.7
4175	201	0.9	4573	330	3.3	4288	400	30.0
4310	400	0.9	4596	330	1.1	4311	400	3.8
4343	u	0.8	4614	420	2.0	4327	400	3.3
4385	u	1.2	4636	420	0.1	4491	500	41.4
4675	510	2.7	4776	430	10.0	4514	500	5.5
4702	321	0.7	4799	430	1.5	4530	500	4.7
4791	222	0.8	4817	520	2.7	4656	510	1.2
4831	312	0.6	4839	520	0.4	4693	**600**	3.6
4864	331	0.7	4939	440	18.4	4716	**600**	0.7
4879	u	0.4	4961	440	2.8	4858	**610**	0.6
4905	421	0.9	4979	530	13.1			
4939	440	0.5	5002	530	2.3			
4966	511	0.6	5019	530	1.7			
4993	322	1.1	5142	540	20.7			
5067	431	1.3	5165	540	3.6			
5085	431	2.4	5182	**630**	5.1			
5108	521	1.0	5205	**630**	1.2			
5155	332	1.8	5304	550	4.6			
5196	422	1.3	5327	550	0.7			
5229	441	1.8	5345	**640**	2.7			
5270	531	1.7	5368	**640**	1.1			
5286	323	0.9	5508	**650**	0.5			
5358	432	4.3						
5398	522	1.4						
5473	**631**	1.3						
5447	404	2.4						
5520	442	3.7						
5561	532	3.9						
5649	433	4.5						
5724	542	4.9						
5740	u	1.1						
5765	**632**	1.5						
5811	443	6.7						
5853	533	4.2						
5887	552	1.0						
5928	**642**	1.0						
5941	434	0.9						
6015	543	8.0						
6056	**633**	1.6						
6103	444	5.6						
6144	534	0.7						
6177	553	1.7						
6218	**643**	2.0						
6306	544	6.8						
6382	**653**	0.3						
6394	u	0.6						
6469	554	1.6						
6510	**644**	0.7						
6597	545	0.8						
6672	**654**	0.4						
6761	555	0.7						

Salt adducts, are assigned and are shown as multiple entries of the same glycan. Hexa-substituted species are in bold and unassigned peaks are shown as u. The glycan content of the glycopeptide is indicated by a three number code, 321, for example, indicating a glycan composition of 3 GalNAc, 2 Gal and 1 NeuAc.

from MALDI-ToF-MS analyses of glycopeptide pools derived from exoglycosidase digestions of this HGP sample. The spectrum obtained following neuraminidase treatment of the pool is shown in Figure 2b and assignments of species 650, 640 and 630 could be made to the peaks observed at *m*/*z* 5508 (calculated 5506), 5345 (calculated 5344) and 5182 (calculated 5182). Other peaks present in this spectrum were readily assigned to tri-, tetra- and penta-substituted glycopeptides (Table 1). Additional β-1→3 galactosidase treatment of this

pool then produced the spectrum as shown in Figure 2c and in this case the *m/z* at 4693, 4491, 4288 and 4086 could be assigned to the species 600 (calculated 4695) 500 (calculated 4491), 400 (calculated 4289) and to 300 (calculated 4086). Minor peaks (<2% of the total) attributable to incomplete de-O-galactosylation, were also present in this spectrum (*m/z* 4858 corresponding to 610 and *m/z* 4656 to 510). Further digestion of this pool with α-*N*-acetylgalactosaminidase resulted, as has been previously reported,[7] in a single species of *m/z* 3476 (Fig. 2d) corresponding to the totally de-glycosylated 33mer peptide demonstrating that a single peptide substituted with a variety of different glycans had been investigated. These results demonstrate that the IgA1 HGP may contain three, four, five or six O-glycan chains. In our studies, over 70 samples from normal controls as well as patients with diseases such as hepatic cirrhosis and glomerulonephritis were similarly found to have the hexa-substitution, implying that this is a widespread feature of IgA1 O-glycosylation both in health and diseased states. Mass spectrometry is not, in general, considered to be a quantitative analytical technique because of variable ionisation abilities of different molecules. However, there are examples where comparisons of the response factors of structurally similar species have been shown to provide good quantitative data.[11,12] For the HGP pools, relative peak intensities/areas have been shown to be sufficiently reproducible to enable estimates of the relative amounts of different glycoforms present.[9] By using the data from Figure 2c, together with similar data obtained from the six other samples investigated in this detail, it is estimated that around 5–10% of human serum IgA1 is hexa-O-glycosylated. The location of the site of substitution of the sixth GalNAc within the peptide chain was then investigated by digesting exoglycosidase-treated samples, separately, with two endoproteases, Glu-C from *Staphylococcus aureus* V8 and proteinase K from *Tritirachium album*. The V8 protease would be expected to cleave the HGP on the carboxyl side of aspartic acid (residue 215, Fig. 1) and the latter protease at several sites, as it possesses a broad specificity towards aliphatic, aromatic and other hydrophobic amino acids. The mass spectrum resulting from V8 digestion of a neuraminidase/galactosidase/*N*-acetylgalactosaminidase-treated sample (see Fig. 2d) contained as expected peaks at *m/z* 962 (assigned to (208)HYTNPSQD(215), calculated *m/z* 961) and 2533 (assigned to (216)VTVPCPVPSTPPTPSPSTPPTPSPS(240) calculated *m/z* 2533). The spectrum obtained from the corresponding neuraminidase/galactosidase-treated sample (see Fig. 2c), which had been similarly-treated with V8 is shown in Figure 3a. Of particular note is the presence in this spectrum of species at *m/z* 963, 3144, 3347, 3550 and 3753 attributable respectively to (208)HYTNPSQD(215) (calculated *m/z* 961) and to the glycopeptides (216)VTVPCPVPSTPPTPSPSTPPTPSPS(240) substituted with 3,4,5 and 6 GalNAc residues (calculated *m/z* 3144, 3347, 3550, 3753). A species with *m/z* 1164, corresponding to the peptide HYTNPSQD plus a single GalNAc residue was not observed suggesting that such a glycopeptide was not present in the digest. Most of the other peaks in Figure 3a, are probably due to peptides derived from autolytic digestion of the protease itself since they were present in control samples containing only the enzyme and were also present in the neuraminidase/galactosidase/*N*-acetylgalactosaminidase-treated sample. These species have not been further investigated. It can be concluded from these results that a sixth glycan may be located at T217 or T233 or S224 or S238 or S240 and not at T210 or S213 (see Fig. 1). It is also possible that the sixth GalNAc is present as the disaccharide GalNAcα-GalNAc at one or more of the five previously identified glycosylated residues. Proteinase K-treated samples were similarly analysed. The spectrum obtained from a neuraminidase/galactosidase-treated sample is shown in Figure 3b. In this instance species at *m/z* 2160, 2362, 2565 and 2769 were observed and these corresponded to the glycopeptides (225)TPPTPSPSTPPTPSPS(240) (or the isobaric (223)PSTPPTPSPSTPPTPS(238)) substituted with 3,4,5 and 6 GalNAc residues (calculated *m/z* 2158, 2361, 2564, 2767). These results indicate that the glycopeptide has been cleaved, at serine 224 (and/or at both valine 222 and serine 238) and consequently that a sixth glycan may be located at either S224 or T233 or S238 or S240 (see Fig. 1). In Figure 3b, as indicated, there were also at least three series of glycopeptides observed at higher mass with each series indicating 3,4,5 and 6 GalNAc substitutions. These correspond to products resulting from proteolytic cleavage at S213, D215 and V216. None of the above glycopeptide species were detected in the neuraminidase/galactosidase/*N*-acetylgalactosaminidase-treated sample that had been treated with proteinase K and those species that were observed were attributed to autolytic peptides (for reasons given above) and to peptides resulting from cleavage at multiple sites within the 33mer peptide. These were not investigated in detail. In summary, the results from these endoprotease digestions suggest that a sixth saccharide is located at one or more of the following residues—224, 233, 238 or 240 or, as mentioned earlier, it is present as GalNAcα-GalNAc. Further investigations will be required to establish the precise location and also to determine its possible role in the pathogenesis of diseases with altered glycosylation. Its presence will however require consideration in future studies concerned with structural/functional properties of IgA1.

1. Experimental

For general methods and further details, see Ref. 9.

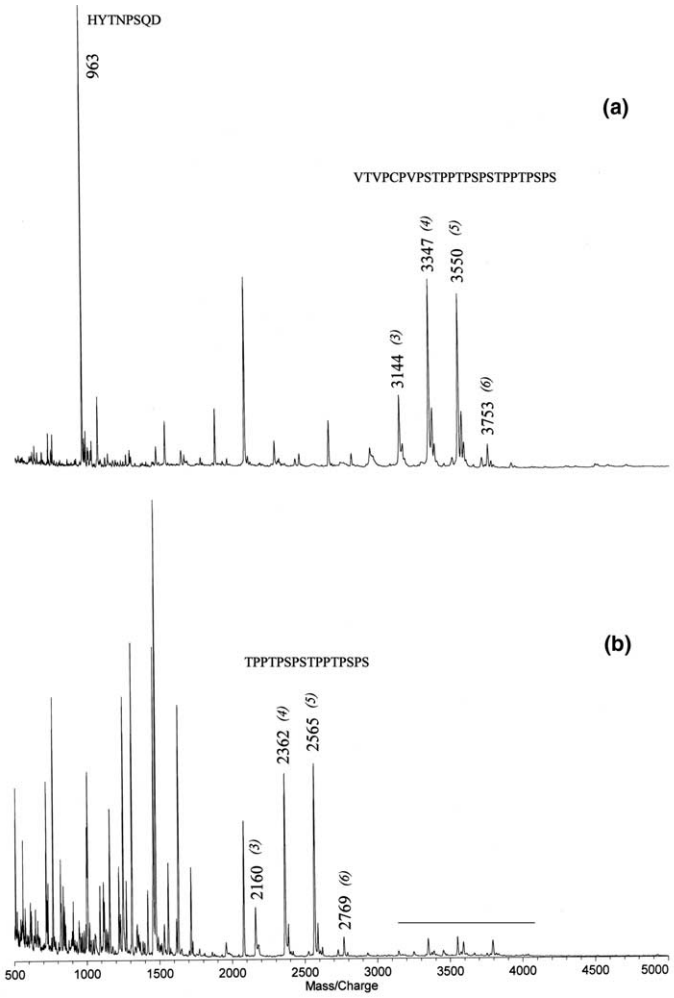

Figure 3. MALDI-ToF-MS of (a) post β-1→3 galactosidase treated sample after digestion with V8 endoprotease; (b) post β-1→3 galactosidase treated pool after digestion with endoproteinase K. *m/z* of (glyco)peptides are shown, and where appropriate, italicised in parenthesis, the number of associated GalNAc residues. Peptide sequences are as indicated. (—) Indicates glycopeptides arising from cleavage at S213, D215 and V216. A different IgA1 sample to that used in producing Figure 2a–d was used for the proteolytic digestion studies. It had a different ratio of 3,4,5 and 6 substitution and this remained consistent throughout the digestion experiments.

1.1. Isolation of IgA1 O-glycopeptides from human serum

This was carried out essentially as previously described.[7,9] Briefly, human serum was precipitated with 50% saturated ammonium sulfate and the precipitate fractionated on Jacalin-sepharose. Following reduction (dithiothreitol), alkylation (4-vinylpyridine) and trypsin digestion, further fractionation on Jacalin-sepharose gave a pool of O-linked glycopeptides–hinge glycopeptides (HGP)—which was then given a final purification using C_{18} reversed phase HPLC.

1.2. De-glycosylation of HGP pool

The HGP pool was sequentially treated at 37 °C with neuraminidase, followed by β-1→3 galactosidase and finally with α-*N*-acetylgalactosaminidase according to the manufacturers instructions. Samples removed after each treatment were stored at −20 °C until required and then desalted using a C_{18} Ziptip® (Millipore) prior to MALDI analysis.

1.3. Protease digestions following digestion with exoglycosidases

The samples following β-1→3 galactosidase and α-*N*-acetylgalactosaminidase treatment (above) were treated with either endoproteinase Glu-C from *S. aureus* V8 (E.C.3.4.21.19) (1 unit) or proteinase K from *T. album* (E.C. 3.4.21.64) (88 munits) at pH 7.5 (50 mM phosphate, 0.1% NaN_3), 37 °C for 17 h. Control samples containing only enzyme were similarly prepared. The digests were desalted using a C_{18} Ziptip® (Millipore) prior to MALDI analysis.

1.4. MALDI-ToF-MS

MALDI-ToF-MS was performed on a Kratos Axima CFR instrument (Kratos Analytical, Ltd, Manchester UK) operating in positive linear mode. 2,4,6-Trihydroxyacetophenone (3 mg/mL in 1:3 CH_3CN: 50 mM aqueous dibasic ammonium citrate) was used as matrix. A glycopeptide pool at a total concentration of around 50–100 pmol/μL (0.5 μL) was applied to the MALDI target plate and whilst still liquid, 0.5 μL of matrix was added and the mixture dried rapidly (<1 min) under vacuum.

Acknowledgements

Use of the Mass Spectrometry facilities in The Medical Biomics Centre, St. George's Hospital Medical School is gratefully acknowledged. Financial support from The Charitable Foundation of Guy's and St. Thomas' Hospitals is gratefully acknowledged. We would like to acknowledge the assistance of Dr. Patrick Corran for purifying the hinge glycopeptides by rpHPLC, Mr. Howard Smith for devising the Java programme facilitating the handling of the data obtained from the mass spectrometer and Mr. Glenn Alder for assistance with preparation of the figures.

References

1. Baenziger, J.; Kornfeld, S. *J. Biol. Chem.* **1974**, *249*, 7260–7269.
2. Field, M. C.; Amatayakul-Chantler, S.; Rademacher, T. W.; Rudd, P. M.; Dwek, R. A. *Biochem. J.* **1994**, *299*, 261–275.
3. Baenziger, J.; Kornfeld, S. *J. Biol. Chem.* **1974**, *249*, 7270–7281.
4. Mattu, T.; Pleass, R. J.; Willis, A. C.; Kilian, M.; Wormald, M. R.; Lellouch, A. C.; Rudd, P. M.; Woof, J. M.; Dwek, R. A. *J. Biol. Chem.* **1998**, *273*, 2260–2272.
5. Odani, H.; Hiki, Y.; Takahashi, M.; Nishimoto, A.; Yasuda, Y.; Iwase, H.; Shinzato, T.; Maeda, K. *Biochem. Biophys. Commun.* **2000**, *271*, 268–274.
6. Allen, A. C.; Bailey, E. M.; Brenchley, P. E. C.; Buck, K. S.; Barratt, J.; Feehally *Kidney Int.* **2001**, *60*, 969–973.
7. Iwase, H.; Tanaka, A.; Hiki, Y.; Kokubo, T.; Ishii-Karakasa, I.; Kobayashi, Y.; Hotto, K. *J. Biochem. (Tokyo)* **1996**, *120*, 393–397.
8. Horie, A.; Hiki, Y.; Odani, H.; Yasuda, Y.; Takahashi, M.; Kato, M.; Iwase, H.; Kobayashi, Y.; Naakashima, I.; Maeda, K. *Am. J. Kidney Dis.* **2003**, *42*, 486–496.
9. Pouria, S.; Corran, P. H.; Smith, A. C.; Smith, H. W.; Hendry, B. M.; Challacombe, S. J.; Tarelli, E. *Anal. Biochem.* **2004**, *330*, 257–263.
10. Papac, D. I.; Wong, A.; Jones, A. J. *Anal. Chem.* **1996**, *68*, 3215–3223.
11. Harvey, D. J. *Rapid Commun. Mass Spectrom.* **1993**, *7*, 614–619.
12. Tarelli, E.; Byers, H. L.; Wilson, M.; Roberts, G.; Homer, K. A.; Beighton, D. *Anal. Biochem.* **2000**, *282*, 165–172.