Universiteit
Leiden
The Netherlands

# Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins

Kozak, R.P.

**Citation**
Kozak, R. P. (2017, January 24). *Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins*. Retrieved from https://hdl.handle.net/1887/45434

| | |
|---|---|
| Version: | Not Applicable (or Unknown) |
| License: | Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden |
| Downloaded from: | https://hdl.handle.net/1887/45434 |

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/45434 holds various files of this Leiden University dissertation.

**Author**: Kozak, R.P.
**Title**: Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins
**Issue Date**: 2017-01-24

# Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins

Radoslaw Pawel Kozak

ISBN: 978-94-6295-566-0

© 2016 Radoslaw Pawel Kozak. All rights reserved. No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means without permission of the author or the journals holding the copyrights of the published manuscripts. All published material was reprinted with permission.

The work presented in this thesis was performed at Ludger Ltd, Oxfordshire, UK and Centre for Proteomics and Metabolomics from the Leiden University Medical Centre.

The studies described in this thesis were supported by Ludger Ltd.

The printing of this thesis was financially supported by Ludger Ltd.

Photography: Matthew Doherty (Ludger Ltd)

Cover design: Ewelina Semba

Layout and printing: Proefschriftmaken

# Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 24 januari 2017
klokke 11.15 uur

door

Radoslaw Pawel Kozak
geboren te Wloclawek
op 05 november 1983

| | |
|---|---|
| **Promotores:** | Prof. Dr. A.M. Deelder |
| | Prof. Dr. M. Wuhrer |
| **Co-promotor:** | Dr. D.L. Fernandes |
| **Leden promotiecommissie:** | Prof. Dr. S. Flitsch |
| | (University of Manchester) |
| | Prof. Dr. H.J. Guchelaar |
| | Prof. Dr. H.S. Overkleeft |
| | Prof. Dr. H. P. Spaink |
| | Dr. C.Balog |

# Table of Content


## Chapter 1

*General introduction* **7**

## Chapter 2

*Protein O-glycosylation analysis* **21**

## Chapter 3

*Suppression of peeling during the release of O-glycans by hydrazinolysis.* **75**

## Chapter 4

*Improved nonreductive O-glycan release by hydrazinolysis with ethylenediaminetetraacetic addition.* **103**

## Chapter 5

*Comparison of procainamide and 2-aminobenzamide labelling for profiling and identification of glycans by liquid chromatography with fluorescence detection coupled to electrospray ionization-mass spectrometry.* **127**

## Chapter 6

*Variation of human salivary O-glycome.* **139**

## Chapter 7

*General discussion.* **165**

## Addendum

*English Summary* **178**

*Dutch Summary* **182**

*Curriculum Vitae* **191**

*List of publications* **192**

*Acknowledgements* **195**

# Chapter 1

## General introduction

## Glycosylation

Glycosylation is the enzymatic process in which oligosaccharides (known as glycans) are added to proteins, lipids or other organic molecules. Protein glycosylation is the most common post-translational modification found in all domains of life. It has been estimated that more than 50% of human proteins are glycosylated.[1] Proteins are made up of combinations of 20 different amino acids and glycans are built from several monosaccharides including: galactose (Gal), glucose (Glc), fucose (Fuc), mannose (Man), *N*-acetylgalactosamine (GalNAc), *N*-acetylglucosamine (GlcNAc), and *N*-acetylneuraminic acid (NeuAc).[2] In this context protein glycosylation can display structural heterogeneity with respect and with respect to the glycan's structure. There are two main types of protein glycosylation, N-linked glycosylation and O-linked glycosylation (Figure 1).

The N-linked glycans are covalently attached to the amino group in the asparagine side chain18 in the sequence Asn-X-Ser/Thr (where X is any amino acid except proline). The glycan is later modified in the ER and Golgi apparatus, creating the three types of N-glycans found in mature glycoproteins: high mannose, complex and hybrid with the core structure $Man_3GlcNAc_2$.[3,4]

The O-linked glycans are covalently attached to the hydroxyl group of serine or threonine within a peptide backbone. A range of monosaccharides linked to serine or threonine have been found (e.g. *N*-acetylglucosamine, *N*-acetylgalactosamine, xylose, mannose and fucose). Extended structures from a core *N*-acetyloglucosamine also can be found and these are called mucin type O-glycans.[5-7] The focus of this thesis is on the development of analytical methods for analysis of mucin type O-glycosylation.

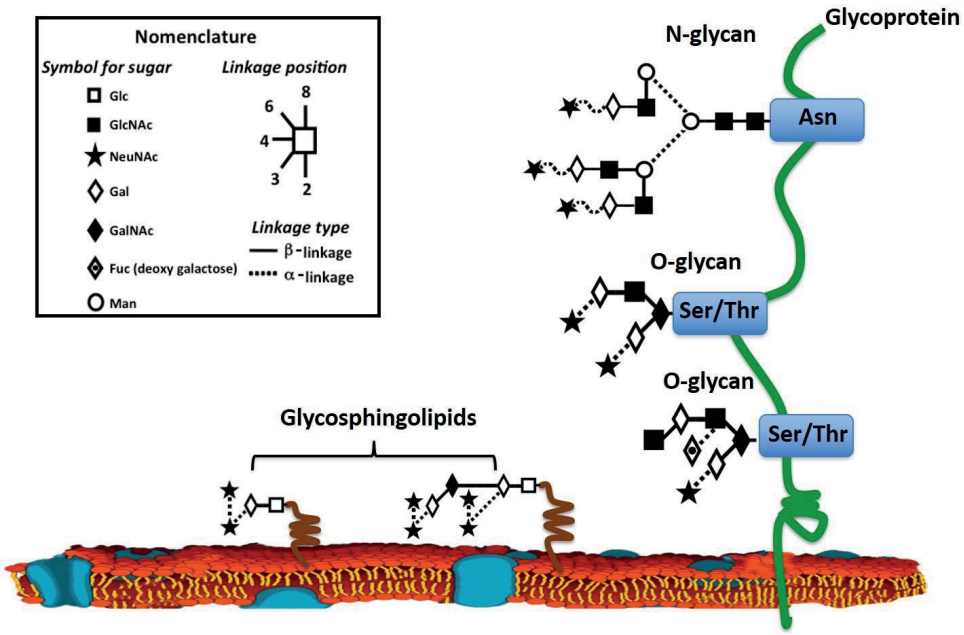



**Figure 1.** Types of glycosylation. Glycans are schematically represented according to Harvey et al.[8]

## Biopharmaceuticals

Over the last decade the rate of approved protein therapeutics (biopharmaceuticals or biologic drugs) in the United States and European Union has been rapidly growing.[9] The cost of developing and bringing biopharmaceuticals to the market is much higher (when compared to classic small molecule drugs), offering increased profit margins. The biopharmaceutical market is currently estimated at US$199.7 billion globally and it is expected to reach US$ 497.9 billion by 2020, growing with an approximate annual rate of 13.5 %.[10]

Biotherapeutic development has been focused on treating a wide variety of chronic diseases including diabetes, cancer, hepatitis C and B and chronic renal failure – as well as less common ones such as hemophilia, Fabry's disease, growth deficiency, multiple sclerosis and Crohn's disease. The appeal of biopharmaceuticals

is that they harness the power of natural therapeutically active biomolecules and often have greater efficacy and specificity than small molecule drugs. The large majority of biologic drugs have been produced in cell culture and they are highly complex glycoproteins composed of various polypeptide domains, some of which are glycosylated.[11] There are two main types of glycosylation present in biopharmaceuticals: N- and O-glycosylation and both can impact safety, efficacy, immunogenicity, solubility, protein folding and half-life of the drug [12,13] (Table 1).

Given the potential impact of glycosylation on the safety and efficacy profiles of biologic drugs biopharma companies must optimise, measure and control the glycosylation patterns of their therapeutics. Unfortunately, there are three aspects that make these tasks very difficult. Firstly glycans are extremely complex molecules and most glycoprotein therapeutics have more than one glycosylation site. Secondly, glycans cause structural and functional diversity of the proteins to which they are attached. Most glycoprotein drugs exist not as a single molecule but as a heterogeneous mixture of glycoforms which share the same protein but which differ in the glycans. Each of these glycoforms might have its own distinct clinical safety and efficacy profile. Thirdly, changes in glycosylation patterns are the major source of batch to batch variability for glycoprotein therapeutics.[13] To overcome these challenges biopharma companies have to employ analytical technologies for detailed characterisation of glycosylation during drug design and biomanufacturing. In parallel with this is a drive from health agencies such as the US Food and Drug Administration (FDA), the European Medicines Evaluation Agency (EMEA), and the International Conference on Harmonization (ICH) which have produced guidance

| Generic Designation | Therapeutic Protein | Expression System | Glycosylation |
|---|---|---|---|
| Algulcosidase alfa | Acid glucosidase | CHO cells | N-links[14] |
| Anti-hemophilic factor | Factor VIII | BHK & CHO cells | N-links & O-links[15] |
| Chorionic Gonadotropin alfa | Chorionic Gonadotropin | CHO cells | N-links & O-links[16,17] |
| Darbepoietin alfa | Engineered erythropoietin with 2 extraN-glycans (long-acting) | CHO cells | N-links & O-links[16,17] |
| Epoetin alfa | Erythropoietin | CHO cells | N-links & O-links[18,19] |
| Etanercept | A dimeric fusion protein consisting of the extracellular ligandbinding portion of the human 75 kilodalton (p75) tumor necrosis factor receptor (TNFR) linked to the Fc portion of human IgG1 | CHO cells | N-links & O-links[18-20] |
| Factor VIIa | Factor VIIa | BHK cells | N-links & O-links[21] |
| Factor IX | Factor IX | CHO cells | N-links & O-links[22,23] |
| Hepatitis B vaccine | Surface antigen of Hepatitis B | *S. cerevisiae* | N-links & O-links[24] |
| Hepatitis B vaccine | S, pre-S and pre-S2 | Murine cell line | N-links & O-links[24] |
| Lenograstim | Granulocyte colony stimulating factor (G-CSF) | CHO cells | O-links[25] |
| Thyrotropin alfa | Thyroid stimulating hormone | CHO cells | N-links[26,27] |

**Table 1.** Examples of approved glycosylated therapeutics.

documents relating to regulatory requirements for measurement and control of drug glycosylation.[28,29] The variability of drug glycosylation is a major concern to biopharma companies and regulators as it can cause inconsistencies in the clinical performance of the drug product. The consequences are potential risks to patients being treated with drugs that are either ineffective or too strong. This also exposes companies to litigation for producing unreliable therapeutics. Given this, there is considerable interest in mitigating these risks by using a Quality by Design (QbD) to deal with glycosylation. In practice, this entails identifying, measuring and optimising

Glycosylation Critical Quality Attributes (GCQAs) starting at an early stage in drug development and continuing through the product's life cycle.[29,30] To be successful, QbD of drug glycosylation requires use of a range of orthogonal methods for glycosylation structure analysis. In practice, these are drawn from three main and complementary analytical platforms: high-performance liquid chromatography (HPLC), mass spectrometry (MS) and capillary electrophoresis (CE).

## Biosynthesis and functions of O-linked glycans

Biosynthesis of mucin type O-glycans is very complex and is initiated in the Golgi apparatus by a *N*-acetyl galactosaminyltransferase (UDP-GalNAc) that transfers a *N*-acetylgalactosamine (GalNAc) residue to the side chain of a serine or a threonine in a glycoprotein via α-linkage to form GalNAcα1- Ser/Thr, which is also known as the Tn antigen.[31,32] Stepwise enzymatic elongation by specific transferases produces several core structures, which are further elongated or modified by sialylation, sulfatation, acetylation, fucosylation, and polylactosamine-extension.[33,34]

Synthesis of Tn antigen is followed by transfer of galactose (Gal), *N*-acetylglucosamine (GlcNAc) or *N*-acetylgalactosamine (GalNAc) to the Tn antigen to form eight core O-glycan structures (Figure 2). In addition core 1–4 structures are extended to form various structures including polyLacNAc chains[35], Lewis antigens[36], and various blood group antigens including the well-known ABO blood groups.[36]

The glycans associated with proteins modulate protein function and signaling. They can also alter the dynamics of glycoprotein endocytosis and cell surface half-life through binding to multivalent lectins.[37] Glycans play important roles in biological processes such as protein secretion (as they influence protein folding)[38], protein

clearance [39], receptor binding and activation [40] and cell adhesion.[41] Glycans also contribute to the regulation of cytosolic and nuclear functions, immune surveillance, inflammatory reactions, auto-immunity, hormone action and tumour metastasis.[42,43] Specific glycan patterns can change when the physiological state changes – for example, during disease or aging.[44,45] Protein glycosylation also plays very important role during absorption, distribution, metabolism and excretion of a drug in an organism.[46,47]

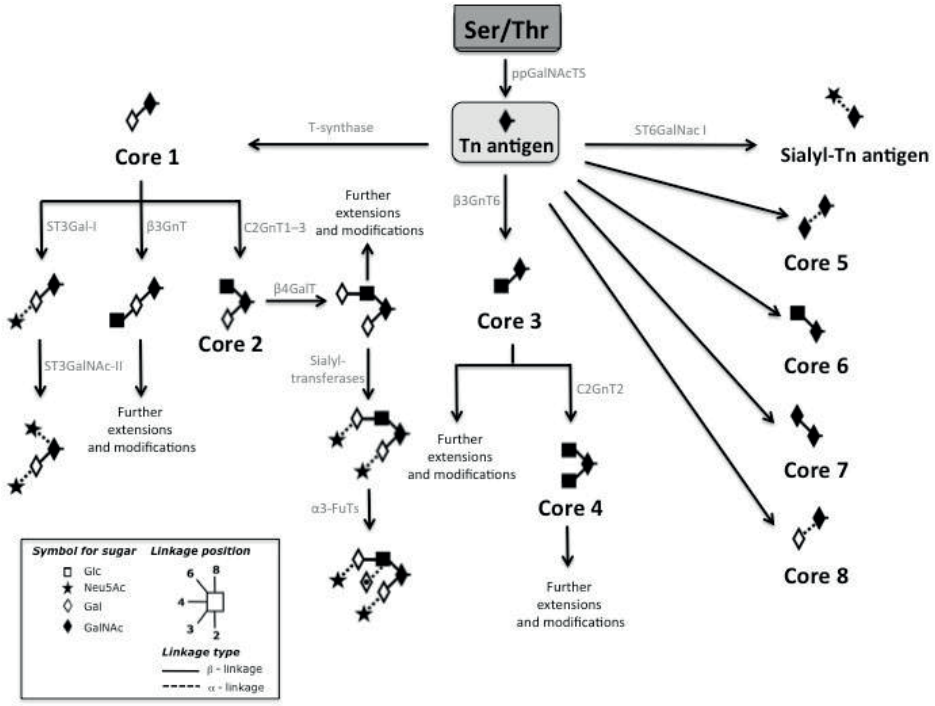



**Figure 2.** O-glycan biosynthesis pathway.

## Functions of O-linked glylcans

During the last decade or so, many aspects of the biological roles of mucin type O-glycans have been reported. Functions of O-glycans are very diverse and include among others: (i) they maintain the highly extended and rigid structure of mucins, as

clustered O-glycans induce the peptide core to adopt a stiff and extended conformation that prevents folding into a globular structure;[48-50] (ii) they contribute to preventing the desiccation of the ocular surface by providing a hydrophilic character to mucins;[51] (iii) they play a role in the processing and expression of glycoproteins; (iv) they are involved in recognition in different processes (e.g. selecting binding in leukocyte circulation, fertilization, glycoprotein clearance);[50,52] (v) they are also known to restrict adenoviral vector access to apical receptors expressed in respiratory epithelium.[53]

Changes in O-glycosylation have been associated with many diseases such as inflammatory bowel disease, cystic fibrosis, IgA nephropathy and cancer. Inflammatory bowel diseases and cystic fibrosis are characterized by alterations in sulfation and sialylation of terminal O-glycan chains, which may result in a detrimental effect on the physico-chemical properties of the mucus, its protective barrier function, which would in turn impact bacterial colonization.[54,55] IgA nephropathy is an immune disorder characterized by reduced galactosylation of the Tn-antigen in IgA1.[56-58]

Mucins produced in cancer cells have an altered expression of glycosyltranferases that results in short and highly sialylated *O*-glycan chains, such as the Tn and sialyl-Tn antigens. Changes in O-glycan structures in cancer have many consequences for the function of epithelial cells, altering their antigenic and adhesive properties, as well as their potential to invade and metastasize.[59-62] Some of these functional properties of mucin type O-glycans might be used to design new biological drugs.

Aspects of glycosylation analysis will be extensively described in Chapter 2.

## Scope of the thesis

Over the last decade the number of techniques for O-glycan analysis have been developed (described in Chapter 2), however release of O-glycans remains very challenging. The main reason for this is that there is no enzyme available for universal O-glycan release from proteins. Therefore, chemical release is the most effective method for release full range of O-glycans. Unfortunately all of the O-glycan release methods show a side reaction known as “peeling”, which is a stepwise degradation of the polysaccharide starting at the reducing end and removing one sugar residue at a time. Peeling is well known problem when performing the O-glycan release and often results in poor repeatability with variable amounts of the small peeled glycans. This is a major problem for comparability studies or quality control. The scope of this thesis was to develop or improve methods for the release of O-glycans with free reducing termini with high yield and limited degradation for the more detailed characterization of biological samples and biopharmaceuticals. In Chapters 3 and 4 the development of sample preparation techniques for O-glycan release (in their nonreduced form) by hydrazinolysis are described. In Chapter 5, the use of procainamide for fluorescent labelling of glycans is introduced. The procainamide labelled glycans gave higher fluorescent response and significantly higher ion intensity in positive ESI-MS when compared to most widely used 2-AB. Chapter 6 describes the application of the improved protocols for O-glycan release by hydrazinolysis combined with developed procainamide labelling system followed by HILIC-UHPLC-ESI-MS/MS analysis for evaluation of O-glycosylation changes in human saliva. The closing chapter of this thesis, Chapter 7, contains a general discussion which places the individual chapters into context, is included.

## Reference List:


(1) Shental-Bechor, D.; Levy, Y. *Curr. Opin. Struct. Biol* **2009**, *19*, 524-533.

(2) Moremen, K. W.; Tiemeyer, M.; Nairn, A. V. *Nat. Rev. Mol. Cell Biol* **2012**, *13*, 448-462.

(3) Kornfeld, R.; Kornfeld, S. *Annu. Rev. Biochem* **1985**, *54*, 631-664.

(4) Nilsson, I. M.; von, H. G. *J. Biol. Chem* **1993**, *268*, 5798-5801.

(5) I., B.; H., S.; P., S. In *Essentials of Glycobiology. 2nd edition*; Cold Spring Harbor Laboratory Press, 2009.

(6) Spiro, R. G. *Glycobiology* **2002**, *12*, 43R-56R.

(7) Tran, D. T.; Ten Hagen, K. G. *J. Biol. Chem* **2013**, *288*, 6921-6929.

(8) Harvey, D. J.; Merry, A. H.; Royle, L.; Campbell, M. P.; Dwek, R. A.; Rudd, P. M. *Proteomics* **2009**, *9*, 3796-3801.

(9) Walsh, G. *Nat. Biotechnol* **2010**, *28*, 917-924.

(10) *Biopharmaceuticals - A Global Market Overview*, 2013.

(11) Higgins, E. *Glycoconj. J* **2010**, *27*, 211-225.

(12) Hossler, P.; Khattak, S. F.; Li, Z. J. *Glycobiology* **2009**, *19*, 936-949.

(13) Dotz, V.; Haselberg, V.; Shubhakar, A.; Kozak, R. P.; Falck, D.; Rombouts, Y.; Reusch, D.; Somsen, G. W.; Fernandes, D. L.; Wuhrer, M. *Trends in Analytical Chemistry* **2015**, *73*, 1-9.

(14) Jongen, S. P.; Gerwig, G. J.; Leeflang, B. R.; Koles, K.; Mannesse, M. L.; van Berkel, P. H.; Pieper, F. R.; Kroos, M. A.; Reuser, A. J.; Zhou, Q.; Jin, X.; Zhang, K.; Edmunds, T.; Kamerling, J. P. *Glycobiology* **2007**, *17*, 600-619.

(15) Hironaka, T.; Furukawa, K.; Esmon, P. C.; Fournel, M. A.; Sawada, S.; Kato, M.; Minaga, T.; Kobata, A. *J. Biol. Chem* **1992**, *267*, 8012-8020.

(16) Kessler, M. J.; Mise, T.; Ghai, R. D.; Bahl, O. P. *J. Biol. Chem* **1979**, *254*, 7909-7914.

(17) Weisshaar, G.; Hiyama, J.; Renwick, A. G. *Glycobiology* **1991**, *1*, 393-404.

(18) Takeuchi, M.; Kobata, A. *Glycobiology* **1991**, *1*, 337-346.

(19) Lai, P. H.; Everett, R.; Wang, F. F.; Arakawa, T.; Goldwasser, E. *J. Biol. Chem* **1986**, *261*, 3116-3121.

(20) Hokke, C. H.; Bergwerff, A. A.; Van Dedem, G. W.; Kamerling, J. P.; Vliegenthart, J. F. *Eur. J. Biochem* **1995**, *228*, 981-1008.

(21) Pennica, D.; Lam, V. T.; Weber, R. F.; Kohr, W. J.; Basa, L. J.; Spellman, M. W.; Ashkenazi, A.; Shire, S. J.; Goeddel, D. V. *Biochemistry* **1993**, *32*, 3131-3138.
(22) Fischer, B.; Mitterer, A.; Dorner, F.; Eibl, J. *J. Thromb. Thrombolysis* **1996**, *3*, 57-62.
(23) Agarwala, K. L.; Kawabata, S.; Takao, T.; Murata, H.; Shimonishi, Y.; Nishimura, H.; Iwanaga, S. *Biochemistry* **1994**, *33*, 5167-5171.
(24) Schmitt, S.; Glebe, D.; Tolle, T. K.; Lochnit, G.; Linder, D.; Geyer, R.; Gerlich, W. H. *J. Gen. Virol* **2004**, *85*, 2045-2053.
(25) Kubota, N.; Orita, T.; Hattori, K.; Oh-eda, M.; Ochi, N.; Yamazaki, T. *J. Biochem* **1990**, *107*, 486-492.
(26) Zhou, Q.; Park, S. H.; Boucher, S.; Higgins, E.; Lee, K.; Edmunds, T. *Anal. Biochem* **2004**, *335*, 10-16.
(27) Cole, E. S.; Lee, K.; Lauziere, K.; Kelton, C.; Chappel, S.; Weintraub, B.; Ferrara, D.; Peterson, P.; Bernasconi, R.; Edmunds, T.; . *Biotechnology (N. Y. )* **1993**, *11*, 1014-1024.
(28) Endo, T. *Biol. Pharm. Bull* **2009**, *32*, 765-766.
(29) Schiestl, M.; Stangler, T.; Torella, C.; Cepeljnik, T.; Toll, H.; Grau, R. *Nat. Biotechnol* **2011**, *29*, 310-312.
(30) Fernandes, D. L. In *Biopharmaceutical Production Technology, Volume 1 & Volume 2 (ed G. Subramanian)*; Wiley-VCH Verlag GmbH & Co.KGaA,Weinheim,Germany., 2012.
(31) Ju, T.; Wang, Y.; Aryal, R. P.; Lehoux, S. D.; Ding, X.; Kudelka, M. R.; Cutler, C.; Zeng, J.; Wang, J.; Sun, X.; Heimburg-Molinaro, J.; Smith, D. F.; Cummings, R. D. *Proteomics. Clin. Appl* **2013**, *7*, 618-631.
(32) Ju, T.; Otto, V. I.; Cummings, R. D. *Angew. Chem. Int. Ed Engl* **2011**, *50*, 1770-1791.
(33) Hounsell, E. F.; Davies, M. J.; Renouf, D. V. *Glycoconj. J* **1996**, *13*, 19-26.
(34) Yamashita, Y.; Chung, Y. S.; Horie, R.; Kannagi, R.; Sowa, M. *J. Natl. Cancer Inst* **1995**, *87*, 441-446.
(35) Liu, J.; Jin, C.; Cherian, R. M.; Karlsson, N. G.; Holgersson, J. *J. Biotechnol* **2015**, *199*, 77-89.
(36) Stanley, P.; Cummings, R. D. **2009**.
(37) Dennis, J. W.; Lau, K. S.; Demetriou, M.; Nabi, I. R. *Traffic* **2009**, *10*, 1569-1578.

(38) Hoseki, J.; Ushioda, R.; Nagata, K. *J. Biochem* **2010**, *147*, 19-25.
(39) Fukuda, M. N.; Sasaki, H.; Lopez, L.; Fukuda, M. *Blood* **1989**, *73*, 84-89.
(40) Marth, J. D.; Grewal, P. K. *Nat. Rev. Immunol* **2008**, *8*, 874-887.
(41) Gu, J.; Sato, Y.; Kariya, Y.; Isaji, T.; Taniguchi, N.; Fukuda, T. *J. Proteome. Res* **2009**, *8*, 431-435.
(42) Varki, A. *Glycobiology* **1993**, *3*, 97-130.
(43) Hart, G. W.; Slawson, C.; Ramirez-Correa, G.; Lagerlof, O. *Annu. Rev. Biochem* **2011**, *80*, 825-858.
(44) Freeze, H. H.; Aebi, M. *Curr. Opin. Struct. Biol* **2005**, *15*, 490-498.
(45) Ohtsubo, K.; Marth, J. D. *Cell* **2006**, *126*, 855-867.
(46) Sola, R. J.; Griebenow, K. *BioDrugs* **2010**, *24*, 9-21.
(47) Li, H.; d'Anjou, M. *Curr. Opin. Biotechnol* **2009**, *20*, 678-684.
(48) Jentoft, N. *Trends Biochem. Sci* **1990**, *15*, 291-294.
(49) Hanisch, F. G. *Biol. Chem* **2001**, *382*, 143-149.
(50) Van den Steen, P.; Rudd, P. M.; Dwek, R. A.; Opdenakker, G. *Crit Rev. Biochem. Mol. Biol* **1998**, *33*, 151-208.
(51) Argueso, P.; Gipson, I. K. *Exp. Eye Res* **2001**, *73*, 281-289.
(52) Fukuda, M.; Tsuboi, S. *Biochim. Biophys. Acta* **1999**, *1455*, 205-217.
(53) Stonebraker, J. R.; Wagner, D.; Lefensty, R. W.; Burns, K.; Gendler, S. J.; Bergelson, J. M.; Boucher, R. C.; O'Neal, W. K.; Pickles, R. J. *J. Virol* **2004**, *78*, 13755-13768.
(54) Xia, B.; Royall, J. A.; Damera, G.; Sachdev, G. P.; Cummings, R. D. *Glycobiology* **2005**, *15*, 747-775.
(55) Corfield, A. P.; Carroll, D.; Myerscough, N.; Probert, C. S. *Front Biosci* **2001**, *6*, D1321-D1357.
(56) Barratt, J.; Smith, A. C.; Feehally, J. *Nephrology. (Carlton. )* **2007**, *12*, 275-284.
(57) Allen, A. C.; Bailey, E. M.; Barratt, J.; Buck, K. S.; Feehally, J. *J. Am. Soc. Nephrol* **1999**, *10*, 1763-1771.
(58) Smith, A. C.; de Wolff, J. F.; Molyneux, K.; Feehally, J.; Barratt, J. *J. Am. Soc. Nephrol* **2006**, *17*, 1192-1199.
(59) Brockhausen, I. *Biochim. Biophys. Acta* **1999**, *1473*, 67-95.
(60) Brockhausen, I. *EMBO Rep* **2006**, *7*, 599-604.
(61) Tsuboi, S.; Hatakeyama, S.; Ohyama, C.; Fukuda, M. *Trends Mol. Med* **2012**, *18*, 224-232.

(62) Kudelka, M. R.; Ju, T.; Heimburg-Molinaro, J.; Cummings, R. D. *Adv. Cancer Res* **2015**, *126*, 53-135.

# Chapter 2

## Protein O-glycosylation analysis

**Gerhild Zauner[1], Radoslaw P. Kozak[2], Richard A. Gardner [2], Daryl L. Fernandes [2], André M. Deelder[1] and Manfred Wuhrer[1]**

[1]Biomolecular Mass Spectrometry Unit, Department of Parasitology, Leiden University Medical Center, Postbus 9600, 2300 RC, Leiden, The Netherlands

[2]Ludger Ltd., Culham Science Centre, Oxfordshire, OX14 3EB, United Kingdom

Biological Chemistry (2012) 687-708.

## Abstract

This review provides an overview on the methods available for analysis of O-glycosylation. Three major themes are addressed: (i) analysis of released O-glycans including different O-glycan liberation, derivatization and detection methods; (ii) analysis of formerly O-glycosylated peptides yielding information on O-glycan attachment sites; (iii) analysis of O-glycopeptides, representing by far the most informative but also most challenging approach for O-glycan analysis. While there are various techniques available for the identification of O-linked oligosaccharides, the focus here is on MS fragmentation techniques such as collision induced fragmentation (CID), electron capture dissociation (ECD) and electron transfer dissociation (ETD). Finally, (iv) the O-glycan analytical challenges that need to be met will be discussed.

# Introduction

Protein glycosylation is a very heterogeneous group of post-translational modifications that has been found in essentially all living organisms, ranging from eubacteria to eukaryotes.[1]

Two major types of protein-linked oligosaccharides (N- and O-glycans) are differentiated on the basis of biosynthesis and linkage: while protein N-glycosylation is initiated by the *en bloc* transfer of a 14-monosaccharide building block onto asparagine amide groups of the nascent polypeptide forming an N-glycosidic linkage,[2] O-glycans are built up on the protein by the sequential enzymatic transfer of individual monosaccharides starting with the glycosylation of hydroxyl groups of the side chains of various amino acids including serine and threonine. Protein glycosylation is predominantly found on secreted proteins, the extracellular portions of membrane proteins, and the luminal side of organelles such as endoplasmatic reticulum, Golgi apparatus, and lysosomes.[2]

O-glycosylation in mammals represents a very diverse group of modifications that are often classified on the basis of the innermost monosaccharide. In mucin O-glycans this initiating monosaccharide is an alpha-linked GalNAc, and 20 human GalNAc transferases have been described that differ in polypeptide preference and tissue-specific expression.[3,4] Eight different core structures of GalNAc-linked O-glycans are distinguished dependent on the substitution of GalNAc with beta-linked Gal, beta-linked GlcNAc or alpha-linked GalNAc in the 3-position and/or 6-position (Figure 1).[5] Changes in mucin O-glycosylation appear to be a general phenomenon of malignancies[6,7] and the resulting mucin O-glycopeptide motifs of, for example,

MUC1 have been shown to be targeted by specific autoantibody responses in cancer which may be of diagnostic value.[8]

Another modification, O-mannosylation, has been observed on various proteins from brain and skeletal muscles, and defects in the involved transferases are associated with various congenital diseases.[9] Other O-glycans exhibit an innermost fucose, glucose, galactose, xylose or N-acetylglucosamine.[9-11] A very peculiar case of O-glycosylation is the so-called O-linked N-acetylglucosamine (O-GlcNAc). This modification, termed O-GlcNAcylation, is mainly found on intracellular, nucleocytosolic proteins and is involved in modulating cellular signalling pathways.[12] Even though these diverse types of O-glycosylation are very interesting, a closer description of the specific methodological aspects is out of scope for this review and in the following we will here focus on mucin-type O-glycosylation.

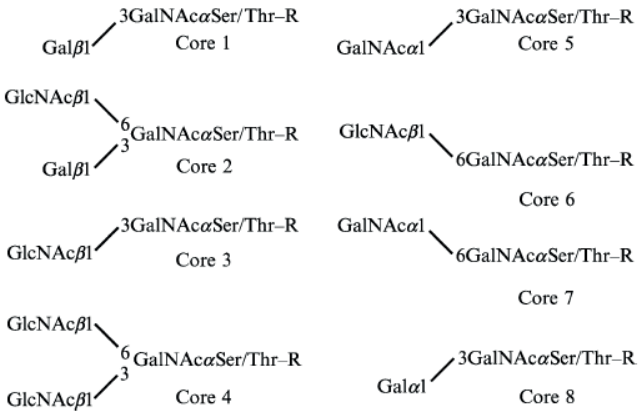



**Figure 1.** Eight types of O-glycan core structures. Taken from Mormann et al.[5] with permission.

A principle challenge in the field of O-glycosylation analysis still comprises finding the most suitable analytical technique that allows the analysis of the desired glycans and their attachment site(s). In general, there are three main ways to analyse glycosylation: (a) analysis and characterization of glycans on intact glycoproteins; (b) analysis and characterization of glycans covalently attached to (glyco)peptides; or (c) structural analysis of released glycans.

The intact glycoprotein, released glycans and generated glycopeptides can be analysed using various analytical platforms such as HPLC profiling, lectin affinity chromatography, mass spectrometric analysis, capillary electrophoresis[13-15] or nuclear magnetic resonance (NMR). Depending on the method used, there are also other steps required to obtain material suitable for analysis; for example in order to perform an HPLC profiling experiment, the glycans need to be released from the protein first and then modified with a fluorescent tag in order to allow optical detection.[13] More recently, Hülsmeier et al.[16] provided a good overview on N- and O-glycosylation analysis methods currently available within the field.[16] The Human Proteome Organisation Human Disease Glycomics/Proteome Initiative initiated an excellent multi-institutional study aiming to compare methodologies currently used for O-glycan analysis. A variety of mass spectrometric and chromatographic procedures representative of current approaches were used.[17] Two general strategies were found to give the most reliable data. These are direct MS analysis of mixtures of permethylated reduced O-glycans in the positive ion mode and analysis of native reduced glycans in the negative ion mode using LC-MS approaches. In addition, mass spectrometric methodologies to analyze O-glycopeptides were also successful. The $^{1}$H NMR studies can provide the entire primary structure of a

released glycan, but the sample needs to be highly purified in large amounts (typically a milligram or more).

Here, we will review current analytical approaches for protein O-glycosylation analysis with a focus on recent methodological advances. Analysis of released glycans, formerly O-glycosylated peptides, and O-glycopeptides in general are covered. Finally, various innovative aspects as well as challenges in the field are discussed.

## Analysis of released glycans

Most methods for protein glycosylation analysis rely on the release of the glycans from the glycoprotein.[18-20] Methods for the release, recovery and analysis of N-glycans are well established with the most common method in use being enzymatic release using N-glycosidase F or N-glycosidase A.[13] However, the release and recovery of O-glycans, particularly intact O-glycans, still remains very challenging. One major reason for this is that the enzymatic release of O-glycans is limited to endo-α-N-acetylgalactosamidase, an O-glycanase with a high specificity that only enables the removal of core 1 disaccharides (Galβ1-3GalNAc) from serine or threonine.[21] Therefore, the current best methods for universal removal of O-glycans from glycoproteins rely on chemical release. Over the last 20 years, several methods for the chemical liberation of O-glycans have been developed including techniques that use either reductive or non-reductive β-elimination.

## Reductive β–elimination

Reductive β-elimination is one of the most popular techniques for obtaining released O-glycans. The most common conditions have been developed by Carlson et al.:[22] the 'Carlson reductive β-elimination method' typically involves the use of 0.05 M sodium hydroxide and 1.0 M sodium borohydride at a temperature of 45°C with an incubation time of 14-16 h followed by solid-phase extraction of released O-glycan alditols.[22-24] Under the alkaline sodium hydroxide conditions, the O-glycans are cleaved from the glycoprotein and the sodium borohydride immediately reduces the terminal monosaccharide of the released sugars to an alditol. Reduction of the terminal sugar by converting it to an alditol reduces or completely stops a side reaction known as 'peeling' which is a side reaction that may be defined as stepwise degradation of the polysaccharide starting at the reducing end and removing one sugar residue at a time.[25-30] Unfortunately, due to the reductive step employed to suppress peeling, the terminal sugar is no longer available for labelling with a fluorophore or chromophore. Therefore, with the preclusion of fluorescence or UV detection, the released O-glycans are predominantly analysed by mass spectrometry or HPAEC-PAD (high pH anion exchange chromatography with pulsed amperometric detection).[31,32] Despite this limitation, reductive β-elimination is still one of the most widely used methods applied by many researchers.[33-37]

Cooke et al.[38] used reductive β-elimination to release mucin-type O-glycans directly from biopsy tissue of rhesus monkeys. The released oligosaccharides were structurally characterized using positive ion mode MALDI -MS and nano-ESI MS for more sensitive detection of acidic oligosaccharides.

Robijn and co-workers used this method for the analysis of the glycosylation of a subset of a *Schistosoma mansoni* egg glycoproteins (SEA-4D12) bound by an anti-carbohydrate mAb 114-4D12 that is used in a diagnostic circulating egg antigen assay. The pool of released O-glycans was analysed by MALDI-TOF-MS.[39]

Analysis of underivatized glycans using negative ion electrospray mass spectrometry can be used for characterizing both negatively charged and neutral oligosaccharides.[33,36,40,41] Isomer separation (on graphitized carbon) and negative-mode ESI MS/MS characterization of underivatized complex O-glycan mixtures from human MUC5B that were released by reductive β-elimination has been described by Karlsson et al.[42] The authors characterized four separated structural isomers of released oligosaccharides with the composition dHex1Hex2HexNAc2HexNAcol1 (Figure 2).[42] Larsson et al.[43] applied reductive β-elimination to study the O-glycosylation of the colon MUC2 mucin from individual biopsies collected from the sigmoid colon during colonoscopy from 25 patients. More than 100 complex O-linked glycans (mostly mono-, di- and trisialylated) were identified and analyzed by graphitized carbon nanoLC-MS/MS.[43]

Lloyd and co-workers[32] used reductive β-elimination to compare the O-linked carbohydrate chains from MUC-1 immunoprecipitated from [3H]GlcN-labelled breast epithelial cell lines (MMSV1–1, MTSV1–7, and HB-2) derived from cells cultured from human milk, with the material obtained from three breast cancer cell lines (MCF-7, BT-20, and T47D) and analysed them by high pH anion chromatography. The data showed that the primary cells had a higher ratio of GlcN/GalN and more complex oligosaccharide profiles than the cancer cell lines.[32]

A significant increase in the sensitivity in the analysis of released O-glycans by MALDI-MS can be achieved by permethylation of the free hydroxyl groups.[34,37,44-46] Permethylating released O-glycans reduces their polarity, increases the stability of glycans (both sialylated and neutral permethylated glycans) for analysis by MS and may provide linkage information upon MS fragmentation or GC-MS analysis of the monosaccharide constituents, for example as partially methylated alditol acetates.[35,47-52]

Hanisch et al.[51] applied permethylation for structural analysis of neutral O-linked glycans from human skim milk mucins. O-glycans were firstly released by reductive β-elimination, then permethylated and subjected to FAB-MS and EI-MS analysis. The major observed structures were Galβ(1-3)[GlcNAcβ(1-6)]GalNAc core 2 type structures.

Thompson and co-workers[41] applied β-elimination to release and characterize O-glycans of salivary mucin. First the oligosaccharide mixture was analysed by NMR, then the oligosaccharides were fractionated into neutral, sialylated and sulphated pools. Neutral and sialylated species were permethylated and analyzed by MALDI-TOF-MS and GC-MS.[41]

Bleckmann et al.[53] applied the reductive β-elimination technique for O-glycosylation analysis of the CD24 glycoprotein from mouse brain. Following the release, the pool of O-glycans was permethylated and successfully analyzed by MALDI-TOF/TOF-MS and ESI-IT-MS/MS (Figure 3). The results indicated the presence of mucin type glycans as well as O-mannosyl glycans which could be distinguished by tandem mass spectrometry.

Faid et al.[34] also used the reductive elimination and permethylation method for the characterization of O-glycan chains in the diagnosis of defects in glycan biosynthesis. O-glycans were released from the glycoproteins of the serum from a patient suffering from CDG type IIa and analyzed by MALDI-TOF-MS and compared with O-glycans released from the glycoproteins of the control serum. No significant difference between O-glycans released from the control serum and from the serum from the patients suffering from CGG type IIa was observed. The major observed O-glycans were: Neu5Ac1Hex1HexNAcol1 and Neu5Ac2Hex1HexNAcol1 consistent with sialylated and disialylated T-antigen structures.

Parry et al.[54] applied this technique for characterization of the O-glycans of CD52 by MALDI-TOF-MS. Presented data indicated that the major O-glycans are core 1 and core 2 mucin type structures with and without sialic acid.

As an alternative to the commonly used sodium hydroxide, dimethylamine has been applied for O-glycan removal by Maniatis et al.[55] The protocol describes the use of dimethylamine in combination with microwave radiation for rapid de-O-glycosylation. This technique was evaluated on bovine fetuin and a porcine stomach mucin. The glycoproteins were incubated with 40% dimethylamine for 70 min and the released and reduced glycans were permethylated and analyzed by MALDI-TOF-MS. The results showed that this procedure 'improved' the yields of O-glycans from both glycoproteins when compared to the 'classic' Carlson reductive β-elimination method.[22,23] A point of criticism in this paper might be the choice of the standard compound (O-GlcNAc-modified peptide) used for method development.[55] O-GlcNAc is known to be much more alkali-labile than O-GalNAc and, therefore, the findings may not directly be transferable to the analysis of mucin-type O-glycosylation.[56]

Although a popular technique, as mentioned before, the analysis of O-glycans released by reductive β-elimination method is largely restricted to mass spectrometric detection.

## Non-reductive β–elimination

Many researchers have looked at developing non-reductive release methods that can produce O-glycans with free reducing termini allowing further derivatization by either a chromophore or a fluorophore for analysis of glycans by HPLC with UV or fluorescence detection.[17] Over the years a number of different reagents for non-reductive release have been developed.

Chai et al.[57] reported the use of 70% (w/v) aqueous ethylamine and Yamada et al.[58] reported the use of lithium hydroxide in an automatic set-up but the overall reaction yields for both methods were low and the amount of peeling products was relatively high.

Huang et al.[59] developed an ammonia based β-elimination procedure. The authors applied a combination of aqueous ammonium hydroxide (for β-elimination) and ammonium carbonate (for converting reducing glycans to glycosylamines as well as glycosylamine carbonates) to prevent O-glycans from 'peeling'. This method is very appealing as both reagents are volatile and easily removed upon evaporation, thus eliminating the need for an elaborative desalting step. The authors reported that the yields of released glycans were quantitative and that peeling was completely prevented. To test this technique the authors used a maltoheptaose standard as a model for both 1-4-linked glycan cores and fetuin N-glycans with a GlcNAcβ1-4GlcNAc core, next to fetuin O-glycans. Released O-glycans were analyzed by

MALDI-MS. No signals originating from the peeling reaction were observable, however the low molecular structures were found to be difficult to detect because they overlap with peaks that are attributed to matrix.[59] A major weakness of this paper might be that inappropriate compounds (β1-4 linked glucose oligomers instead of β1-3 linked glycans, which are much more prone to degradation by peeling) have been used to establish the method. This issue was addressed by Yu et al.[60] who looked into the stability of released O-glycans under non-reducing ammonia based β-elimination conditions demonstrating that the β1-3-linkage is labile whereas the β1-4-linkage is relatively stable under these conditions.[60] These data demonstrate that a complete prevention of peeling under alkali-catalyzed hydrolysis conditions remains difficult. Investigations indicate that great care has to be taken when employing non-reductive alkaline conditions in glycomic analysis and in obtaining glycoprotein glycans for functional studies.[60]

Further modifications to the ammonia based β-elimination procedure were made by Miura et al.[61] They demonstrated that employing ammonium carbamate under the non-reductive conditions can prevent the unfavourable 'peeling' side reaction[61]. The samples were incubated for 20 h at 60°C with ammonium carbamate and then the solution was subjected to glycoblotting with subsequent methylation of sialic acids and tagging of the oligosaccharides by trans-iminization with O-benzylhydroxylamine hydrochloride. The tagged, methyl esterified O-glycans were eluted with water and analyzed by MALDI-TOF-MS.

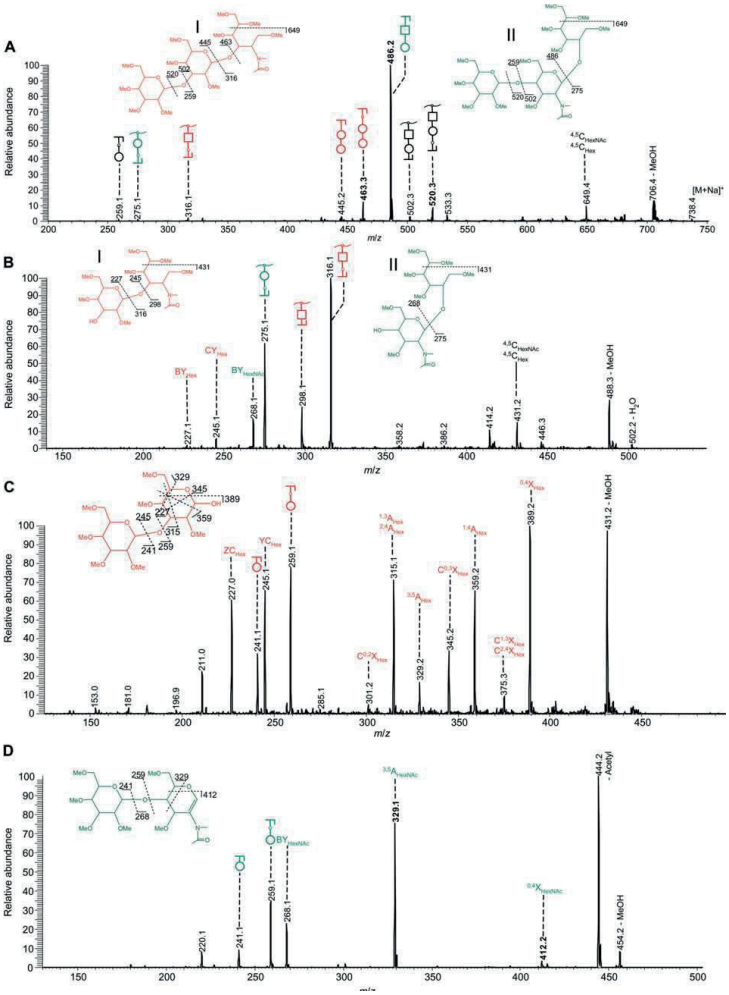



**Figure 2.** ESI-IT-MSn fragmentation spectra of the permethylated, reduced glycan species Hex2HexNAc1. (A) MS2-spectrum of the sodiated pseudomolecular ion $[M+Na]^+$ at m/z 738.4. Ions selected for further fragmentation are given in bold. (B) MS3-spectrum of the Hex-HexNAcol Y ion at m/z 520.2. (C) MS3-spectrum of the Hex-Hex C ion at m/z 463.3. (D) MS3-spectrum of the Hex-HexNAc B-ion at m/z 486.2. Diagnostically relevant fragment ions are illustrated in insets. Possible isomeric structures are given in different colors and key product ions derived thereof are denoted in the respective color. Non-specific product ions are labelled in black. All fragments observed are sodium adducts. Fragments were assigned according to Domon and Costello.[62] They may be achieved by different fragmentation pathways, only one of which is illustrated in the figure. Open circle, hexose; open square, N-acetylhexosamine. Taken from Bleckmann et al.[53] with permission.

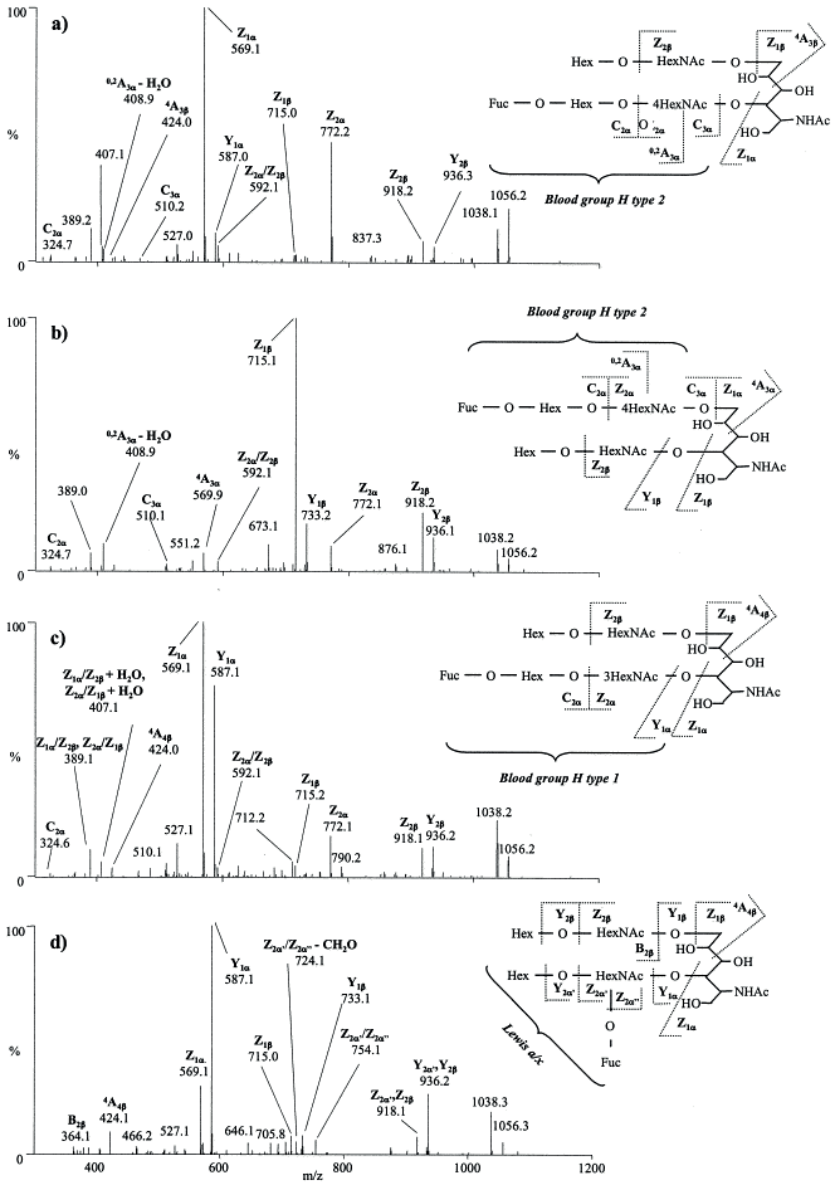



**Figure 3.** LC-MS$^2$ spectra and assigned structures of isomeric Core 4 O-linked oligosaccharides with $[M-H]^-$ ion of m/z 1098 corresponding to dHex1Hex2HexNAc2HexNAcol1 isolated from human MUC5B. Isomers were observed at retention time 19.1 min (a), 18.1 min (b), 16.2 min (c), and 14.9 min (d). Taken from Karlsson et al.[42] with permission.

Permethylation of released O-glycans for analysis by mass spectrometry is a technique used by many researchers. Goetz et al.[63] realised that the alkaline conditions used for permethylation were similar to the conditions used for the release of O-glycans and therefore developed a combined method for the release and permethylation of O-glycans. This method is a combination of an enzymatic digestion and a chemical O-glycan release using a solid-phase permethylation procedure. The protocol consists of incubation of the samples for 48 hours with pronase followed directly by release and permethylation using a spin column method.[64,65] As the glycans are released, in their free, non-reduced form by sodium hydroxide, methyl iodide is used as the methylating reagent to immediately convert them to their permethylated derivatives, thereby preventing the 'peeling' side reaction and stopping the oligosaccharides from degrading. In this study the O-glycans of bovine serum fetuin, human IgA and BSSL were successfully analyzed by MALDI-TOF-MS.[63] Although the technique provides good release of O-glycans, the detection of lower molecular weight O-glycans by MALDI-TOF-MS proves difficult as these masses overlap with peaks in the lower mass region that are attributed to matrix.

More recently, Zauner et al.[66] (2011), Wang et al.[67] and Furukawa et al.[68] have introduced new, one-pot combined methods, for O-glycan release and labelling. These methods use dimethylamine,[66] ammonia[67] or sodium hydroxide[68] for β-elimination release and 1-phenyl-3-methyl-5-pyrazolone (PMP) for the concomitant labelling via a Michael addition. The method was applied to analyse O-glycans from bovine submaxillary gland mucin and bile salt-stimulated lipase (BSSL) by reverse phase nanoLC-ESI-MS/MS. The data shows that the technique is quick and provides good release and labelling with an acceptable amount of peeling products.

Another very interesting study was performed in which chromatographic/tagging methodologies used by three laboratories for the analysis of O-glycans released under non-reductive conditions from IgA1 were found to give a diverse body of data. Clearly, peeling of the non-reductive alkaline released O-glycans from glycopeptides exacerbated the frequent occurrence of artifactual chromatographic peaks, severely influencing the reliability of these methodologies for quantitative O-glycomics (Wada et al., 2010a).[17]

## Hydrazinolysis

Hydrazinolysis is a method for producing O-glycans in a non-reduced form in high yield,[69,70] forming an attractive alternative to the non-reducing β-elimination procedures described above. This method is very versatile and can be applied to release N- and/or O-glycans depending on the reaction conditions. The removal of N-glycans from glycoproteins by hydrazinolysis typically requires a temperature of 85°C to 100°C and an incubation time of 5-16 h.[70] For the removal of O-glycans a lower temperature is required (60°C, 6 h).[69] The mechanism by which hydrazine removes O-glycans, is not fully understood. It is generally accepted that it proceeds via an initial β-elimination reaction followed by reaction with hydrazine to form hydrazone derivatives. Under the release conditions, the N-acetyl groups of the amino sugars are also removed, therefore it is necessary to perform a re-N-acetylation step with acetic anhydride. As with all the non-reductive β-elimination techniques, the major concern of this method is the occurrence of undesirable peeling. A number of researchers have investigated how to reduce peeling under the hydrazinolysis conditions. In order to try and prevent this degradation, Cooper et al. formed hydrazone intermediates using 50% aqueous hydrazine but the results show

that the degradative peeling had not been significantly reduced.[24] More recent studies have shown that the degree of peeling can be greatly reduced by removal of water from the sample before hydrazinolysis[69] or by buffer exchanging the sample prior to hydrazinolysis with a solution of 0.1% trifluoroacetic acid (TFA) or a low molarity solution of ethylenediaminetetraacetic acid (EDTA).[71] Further experiments showed that samples that were contaminated with $CaCl_2$ and subsequently washed with a solution of EDTA showed a lower amount of peeling and an increased yield of released O-glycans. This data demonstrates the effect of salt removal on the suppression of peeling.

Determination of glycan structures from analysing MS, HPLC or NMR data is one of the major challenges in glycomics studies. Several different types of software tools have been developed for glycomics including GlycoWorkbench,[72] GlycoPeakfinder[73] and GlycoBase.[74] GlycoWorkbench is a tool which offers a graphical interface comprising a tool for building glycan schemes, an exhaustive collection of fragmentation types and a broad list of annotation options. The software is publicly available to download from: http://www.glycoworkbench.org/. GlycoPeakfinder is a web based software system allowing determination of the composition of glycan-derived MS signals independently from the source of spectral data and the fragmentation technique. GlycoBase is a database with MS and HILIC data for over four hundred 2-AB labelled N-linked glycans and seventy five 2-AB labelled O-glycan structures. A number of databases for analysing NMR data are also currently available[75,76] but data analysis still is demanding.

## Analysis of formerly O-glycosylated peptides

In most cases the structures of the oligosaccharides are analyzed after releasing O-glycans from peptides by β-elimination. However, there is information that can be obtained from the peptide portion after β-elimination which should not be neglected: whilst no information about the glycan heterogeneity at the different sites can be obtained, the glycosylation sites may be determined.[77]

Initial experiments were performed by Rademaker et al.[78] employing reductive β-elimination with NaOH and $NaBD_4$. This resulted in the released, deutero-labelled O-glycan alditols as well as peptides with a deutero-labelled former O-glycosylation site: for example, in the case of an O-glycosylated threonine, β-elimination of the O-glycan converts the threonine residue into dehydroamino-2-butyric acid which is the reduced to 2-aminobutyric acid deuterated in the C2-position. O-glycosylation sites were elucidated by FAB-MS analysis of the mass-tagged, formerly O-glycosylated peptides providing y- and b-ion series.[78]

A difficulty associated with the β-elimination release of O-glycans from peptides is that the reductant used to prevent peeling of the released O-glycans may degrade the peptide backbone of certain peptides.[79] This was addressed by introducing a β-elimination/Michael addition protocol in which the former site of O-glycan attachment is marked by a tag allowing a mass spectrometric identification of the site of glycosylation.[79] The performance of sodium hydroxide ammonium hydroxide as releasing agents was compared: it was shown that the use of $NH_4OH$ greatly enhances the sensitivity of the β-elimination procedure, as well as obviating the need for sample clean-up. Importantly, the addition of ammonia to the double

bond generated upon glycan elimination was observed, indicating that ammonia had a double function as a basic β-elimination reagent and tagging reagent. The resulting amine-tagged serine and threonine derivatives facilitated the identification of formerly O-glycosylated residues by ESI-quadrupole-CID-MS/MS.[79]

Other researchers have tested other basic chemicals as β-elimination reagents: Hanisch et al.[80] and Mirgorodskaya et al.[81] reported on β-elimination using methylamine as the releasing agent and at the same time as the reagent to modify the formerly glycosylated serine or threonine residue within the peptide prior to MALDI-TOF-MS or ESI-ion trap-MS mapping of the sites.[80,81] Zheng et al.[82] compared the use of various amines (dimethylamine, methylamine and ammonia solution) for the release of O-glycans from glycopeptides via β-elimination.[82] The highest reaction yields were observed when dimethylamine was used to convert the Thr residue of the glycopeptides into its β–elimination/addition derivative with a mass shift of −18 Da or +27Da (addition of dimethylamine) compared to the non-glycosylated Thr. Although the authors successfully identified glyosylation sites, peptide analysis did not give information on the glycan structures.[82]

In two earlier studies removal of O-GlcNAc from O-glycopeptides was described by mild β-elimination followed by Michael addition of dithiothreitol (DTT), deuterated DTT or biotin pentylamine (BAP) to tag O-GlcNAc sites. After chromatographic enrichment of the modified peptides, the former O-GlcNAc sites (as well as phosphorylation sites) could be identified by LC-MS/MS.[83,84]

In a more recent paper, the addition of other reagents during β-elimination, such as 1-phenyl-3-methyl-5-pyrazolone (PMP) in combination with DTT, was tested. It was found that not only were the O-glycans labelled but the formerly glycosylated

peptide portions were also tagged after this reaction, as could be shown by MALDI-TOF-MS. Therefore, after one reaction both the released/labelled O-glycans and the remaining tagged peptide portion can be analysed. This is another demonstration of the feasibility of being able to use the peptide portion for site specific analysis after β-elimination.[68]

Czeszak et al.[85] were likewise interested in analyzing O-glycosylation sites and aimed for a more potent way to chemically modify peptides as they were deglycosylated. They used β-elimination/addition reaction conditions, consisting of dimethylamine and ethanethiol, on formerly glycosylated sites to leave efficiently deglycosylated peptides which were chemically modified to mark the O-glycan attachment sites. After N-terminal derivatization by a phosphonium group, peptide sequencing was carried out by nano-ESI tandem mass spectrometry. [77,85] Depending on the research question, the analysis of deglycosylated peptides can be of value to determine the sites of O-glycosylation, even though the glycan structure and heterogeneity information is lost. [77]

## Analysis of O-glycopeptides

By far the most efficient, but obviously analytically challenging, approach to analyse the O-glycan structures plus the attachment sites to the protein is performing O-glycopeptide analysis. As potential O-glycosylation sites are not indicated by specific consensus motifs, as is the case with N-glycosylation, direct analysis of the glycopeptides is difficult even if using newly developed and improved techniques.[86-88] We will describe the major mass spectrometry techniques currently used for the characterization of O-glycopeptides including CID fragmentation and ECD and ETD fragmentation. Generation of O-glycopeptides by proteolytic cleavage procedures employing specific proteinases and unspecific proteinases will be addressed. The study of synthetic model O-glycopeptides will be presented and examples of naturally occurring O-glycopeptides will be given.

### CID fragmentation of glycopeptides

High energy CID by FAB-MS has been established for characterizing glycopeptides containing O-linked mono- and disaccharides. Via peptide backbone cleavages both the peptide sequence and the site of attachment of the sugar moiety were obtained from a single high-energy CID spectrum.[89-91]

With the advent of ESI and MALDI ionization, glycopeptide analysis by FAB-MS became gradually replaced by these soft ionization modes: the fragmentation of electrospray-ionized glycopeptides may be performed by CID using a variety of instruments and experimental setups. ESI-ion trap-CID tandem mass spectrometry (MS/MS) of glycopeptides results mainly in fragments arising from glycosidic bond cleavage allowing the determination of the glycan structure with the peptide moieties

staying largely intact.[92-94] When multiple sugar units are attached to a peptide, ion series (appearing often in different charge states) are measured due to sequential carbohydrate losses.[92] Under ion trap-CID conditions O-linked glycopeptides often undergo gas-phase deglycosylation that eliminates the carbohydrate completely, leaving the unmodified, and thus, ‘‘unlabelled’’ Ser or Thr residue(s) behind.[94,95] In the case of ion trap-CID, if the detection of peptide backbone fragments is desired, there is in many cases the necessity to perform MS3 experiments.[94,96] In Figure 4 an example of a Proteinase K generated O-glycopeptide identification from asialofetuin is demonstrated: the upper spectrum (Figure 4.A.) shows the precursor ion (m/z 1120.5, [M+2H]2+; inset), and below the MS2 spectrum is shown. As only a few b- and y-ions could be identified in the MS2, an MS3 was necessary to identify the peptide moiety 291VVVGPSVV298 (m/z 755.4; [M+H]+; Figure 4.B.).[94]

While most of the ESI-CID-MS/MS analyses of glycopeptides are performed with positive-mode ionization, Deguchi et al.[97] also evaluated negative-mode ion trap-CID. In these spectra, they observed C-ions comprising the entire O-glycan chain, representing a gas-phase release of the reducing-end O-glycan. Comparison of the negative-mode fragmentation pattern (MS3) of the gas phase-released O-glycan with the MS2 fragmentation of infused oligosaccharide standards enables the use of spectral matching for the structural assignment of O-glycan structures.[97]

Next to ESI-IT-MS/MS, ESI-quadrupole-TOF MS/MS has been widely used for CID fragment ion analysis of O-glycopeptides. ESI-quadrupole-CID fragmentation spectra of O-glycopeptides may allow the deduction of the O-glycosylation site(s) based on y-type and/or b-type peptide ions comprising the glycan attachment site. This is only true if the glycan moiety is not completely ‘lost’ during the fragmentation

process. Hence, optimization of the fragmentation energy is critical in ESI-quadrupole-CID analysis of O-glycopeptides,[98] similar to the situation in ESI-quadrupole-CID of N-glycopeptides where low fragmentation energies result in glycan fragmentation whilst high energies mainly provide peptide sequence ions lacking the glycan modification.[99]

As an example, nano-ESI low energy quadrupole CID was used in conjunction with other methods to analyze N- as well as O-glycopeptides derived from the M surface protein of human hepatitis B virus. The O-glycosylation site within the O-glycopeptide could be identified due to the excellent coverage of the (glycosylated) sequence ions that was obtained for this particularly large peptide (29 amino acids long).[100]

Alving et al.[101] performed a comparative study by ESI-MS/MS and postsource decay (PSD)-MALDI-MS to analyse O-glycopeptides, whereas another study focused exclusively on ESI-QTOF-MS analysis.[98] More recently, MALDI low energy CID MS/MS experiments on a QTOF instrument[102] have been found suitable for assigning O-glycans to specific O-glycosylation sites.

## ECD/ ETD fragmentation of glycopeptides

MS methods using ECD and ETD fragmentation of O-glycopeptides simplify the identification and site assignment of fragmentation-prone side-chain modifications tremendously.

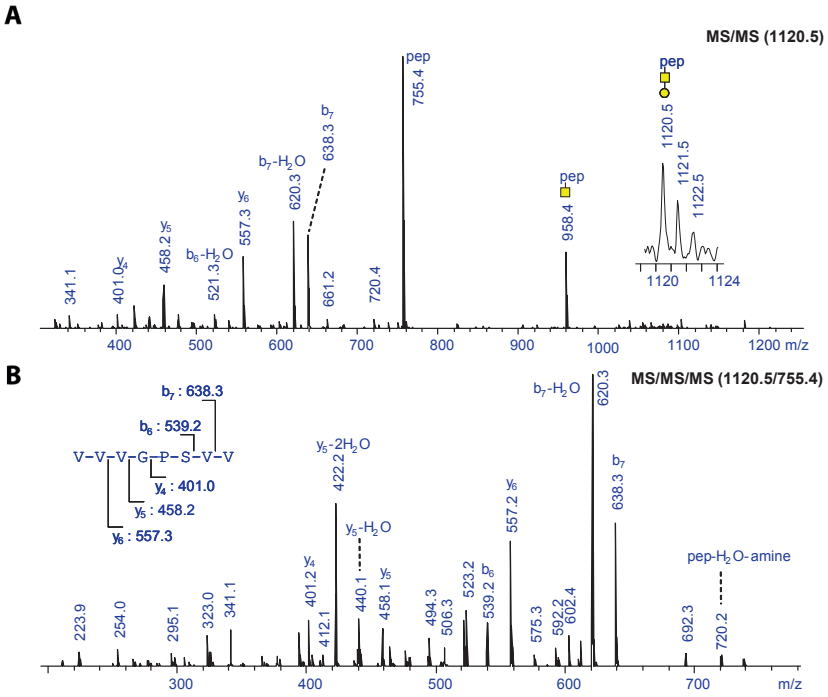



**Figure 4.** Fragment ion spectra of Proteinase K-generated O-glycopeptides from asialofetuin. A fragment ion spectrum (A; MS2) of the protonated precursor (inset) and a corresponding MS3 spectrum (B) are shown. MS2 was performed on glycopeptide V291–V298 carrying a Hex1HexNAc1 O-glycan ($[M+2H]^{2+}$ at m/z 1120.5). The fragment ion at m/z 755.4 ($[M+H]^{+}$) was subjected to MS3. Yellow square, N-acetylgalactosamine; yellow circle, galactose. Taken from Zauner et al.[94] with permission.

In ECD, reduced radical cations $[M+nH]^{(n-1)+}$ are generated by irradiation of multiply charged analyte ions, $[M+nH]^{n+}$, produced by electrospray ionization with low-energy electrons (T≤0.2 eV).[103-106] These odd-electron species mainly dissociate by fast and facile fragmentation of the N–Cα bonds of the peptide chain giving rise to c-type and z-type ions. This process has been characterized as 'non-ergodic', i.e. bond cleavage occurs prior to randomization of the excess energy gained in the charge-recombination event. As a consequence of the non-ergodic nature of the ECD process, the elimination of side chain modifications, such as glycans from the

peptide backbone, cannot compete with the backbone cleavages within the radical cations.[105] Therefore, ECD can be used for an unambiguous assignment of the attachment sites of glycans within the peptide chain.[105,107] In several studies, ECD fragmentation has been proven very useful for the identification of O-glycan structures and their attachment sites within O-glycopeptides.[5,96,105,108-111] For example, Mormann et al.[105] established a protocol for the analysis and characterization of both standard and synthetic O-glycosylated peptides in which a novel pulse sequence improved the efficiency for electron capture dissociation (ECD) of an unmodified Fourier Transform Ion Cyclotron Resonance (FTICR) mass spectrometer by more than an order of magnitude. This approach showed that as well as obtaining the amino acid sequence, information on the attachment site of the labile glycan moiety along with radical-site-induced fragmentation of the glycosidic bonds can be obtained from such an experiment.[105] Renfrow et al.[110,111] made use of ECD and activated ion-electron capture dissociation (AI-ECD) FTICR-MS2 to analyse the O-glycosylation of IgA. Previously established methods relied on the depression or absence of expected signal at particular cycles in the amino acid sequencing, whereas an AI-ECD product ion spectrum directly identifies sites with and without attached glycans.[110] This study led to the first direct identification of multiple sites of O-glycan attachment in IgA1 hinge region by mass spectrometry, thereby enabling future characterization at the molecular level of aberrant glycosylation of IgA1 in diseases. In a follow up study, both FT-ICR MS and liquid chromatography FT-ICR MS were used to analyse the O-glycan heterogeneity of IgA1 proteins. The techniques demonstrated that one could obtain unbiased accurate mass profiles of IgA1 HR glycopeptides from three different IgA1 myeloma proteins.[111] For the first time an individual IgA1 O-glycopeptide species from an IgA1

hinge region preparation that is reproducible for each IgA1 myeloma protein could be identified by AI-ECD fragmentation. Excitation of the precursor ions employing a CO2 laser prior to electron irradiation was found to be valuable for obtaining high-quality ECD spectra of O-glycosylated peptides (in this case IgA O-glycopeptides).[110,111]

In ETD fragmentation, the electron induced fragmentation results in cleavage of the N-Cα bonds of the peptide producing c- and z-type fragment ions.[112,113] Electron transfer triggers a fragmentation almost exclusively along the peptide backbone leaving the glycan mass intact and therefore allows the identification of the peptide sequence. A good quality ETD spectrum permits the identification of the peptide sequence, the glycan mass, and the unambiguous assignment of the modification site (Figure 5).[92]

An advantage of ETD compared to ECD is that the electrons are transferred to the analyte using an electron carrier, which makes the technique compatible with a wider (and more accessible) range of mass spectrometers. Most studies using ETD MS2 as a tool to analyse O-glycopeptides focused on peptides with a single glycan attached (for example (Christiansen et al.,[114] Perdivara et al.[93]) and Table 1), but recently studies have been published on the site-specific characterization of highly clustered glycosylation sites.[113,115]

There are a few disadvantages of ETD such as the requirement of a precursor ion with sufficient charge density for a high quality ETD fragmentation spectrum[92] and the inefficient fragmentation for peptides with a positive net charge <3+.[116] However, most database searching algorithms do not consider fragment ions bearing charge states higher than 2+; they will therefore not assign such ions, and

thus the peptide fragment assignment will be incomplete when highly charged precursors are fragmented.[117]

Another more general problem is the lack of availability of this fragmentation mode on many commonly used MS instruments. Christiansen et al.[114] tried to analyse the mucin-like O-glycosylation domains of a model protein. The protease digests produced many glycopeptides as determined by CID-MS/MS, but ETD fragmentation of these resulted in only a few interpretable spectra, suggesting that the use of ETD for determining the heterogeneous O-glycosylation at specific sites in regions of multiple occupancies is still in its infancy.[114]

In general, it is believed that the combination of CID and ETD fragmentation, along with improved data analysis software should greatly facilitate O-linked glycopeptides analysis in the future.[92]

## Glycopeptides generated by tryptic digestion

The most commonly used enzyme for the proteolytic generation of glycopeptides is trypsin which cleaves the protein at well-defined sites. Specific cleavage makes the analysis of glycopeptides easier as the expected peptide masses can be predicted, if the amino acid sequence of the protein is available.[94]

Steentoft et al.[118] showed with their impressive work that enzymatically glycoengineered human cell lines with simplified O-glycosylation can be used for total O-glycoproteome analysis. Stable 'SimpleCell' lines with homogenous O-glycosylation were generated, which had only truncated GalNAcα or NeuAcα2-6GalNAcα O-glycans attached. Total cell lysates were digested, GalNAc O-glycopeptides isolated and used to analyse the O-glycoproteome by nLC-ESI-LIT-

FTICR-MS. A combination of higher-energy collision-induced dissociation (HCD) and ETD-MS2 mode allowed for peptide sequence analysis with and without retention of

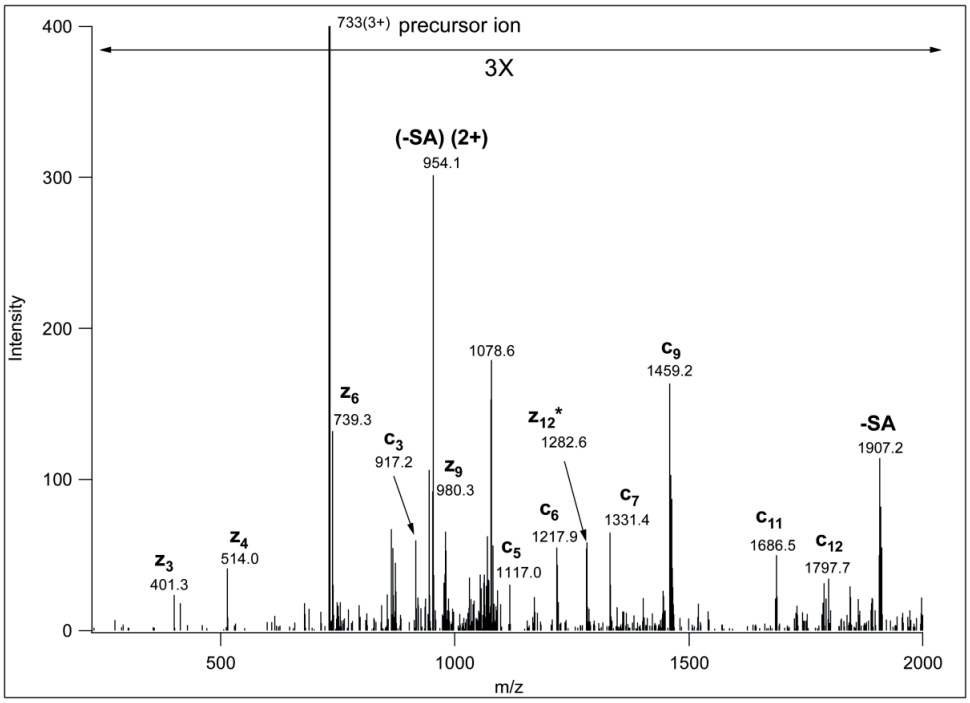



**Figure 5.** ETD spectrum of AAT(GalNAc1Gal1NeuAc1)LSTLAGQPLLER at m/z 733.0369 ($[M+2H]^{3+}$). This spectrum provides sufficient information for confident sequence identification as well as site assignment. Sialic acid loss from charge-reduced versions of the precursor ion was also detected. Fragments labelled with asterisk indicate hydrogen transfer, i.e. z+1 ions. Taken from Darula et al.[92] with permission.

glycan site-specific fragments. The raw data were searched against a database with a precursor mass tolerance of 10 ppm and a fragment ion mass tolerance of 50 millimass units. Methionine oxidation and HexNAc attachment to serine and threonine were used as variable modifications. Additionally, HCD data showing fragment ions at m/z 204.08 were extracted and further processed. The information obtained from ETD as well as HCD spectra was combined for confident glycopeptide

identifications and minimizing the number of false positive hits. In total, >100 O-glycoproteins with >350 O-glycan sites (the great majority previously unidentified), including a GalNAc O-glycan linkage to a tyrosine residue, were identified. By doing the analysis of three human simplified cell lines derived from different organs, the list of known O-glycoproteins could be already doubled while additional sites of O-glycosylation on several previously characterized O-glycoproteins could be identified. This strategy is expected to be amenable to other types of protein O-glycosylation.[118]

In another report, it was shown that a combination of CID and ETD techniques in LC-MS could prove to be a powerful tool for de novo identification of O-glycosylation at unknown modification sites in proteins.[93] More specifically, amyloid precursor protein (APP) was investigated in terms of its glycosylation. There are two previously reported N-glycosylation sites on APP but the authors focused on O-glycosylation plus the identification of their attachment sites: for the first time three threonines were found to be modified with core 1 type O-glycans. This discovery was possible due to combining information obtained from nanoHPLC-MS CID spectra (sugar oxonium ions), and the highly valuable ETD fragmentation pattern of the glycopeptides. In ETD mainly peptide fragmentation occurred with little or no cleavage of the oligosaccharide portion. Clearly, this is also a demonstration that using nanoHPLC in combination with sophisticated fragmentation techniques such as ETD has great potential for mapping unknown glycosylation sites in proteins.[93]

A method to enrich for glycoproteins from proteomic samples was developed recently.[96] In this study N- and O-glycosylation of proteins from human cerebrospinal fluid (CSF) were investigated. For this particular purpose, the sialic acids within the oligosaccharides linked to glycoproteins from CSF were mildly oxidised by periodate,

further captured onto hydrazide beads via the sialic acid residue of the glycans portion, trypsin digested such that only the glycopeptides remain attached to the beads, and finally the captured glycopeptides can be selectively released by acid hydrolysis. After release from the beads, the glycopeptides were analysed by reverse phase (RP)-LC-ESI-FTICR-MS. Low energy CID fragmentation MS2 and MS3 spectra were acquired, resulting in the identification of 36 N-linked and 44 O-linked glycosylation sites on glycoproteins from CSF. In addition, complementary analyses were conducted in which the precursor ions were guided to the ICR cell and fragmented using ECD. This study led to the discovery of five new O-linked glycosylation sites. Glycopeptide fragment spectra were searched against the Mascot database after entering the mass of the peptide portion as a precursor, computed by subtracting the monoisotopic mass of the glycan from the monoisotopic mass of the glycopeptide precursor. The following settings were applied: precursor mass tolerance, 10 ppm; fragment ion mass tolerance, 0.6 Da; enzyme, trypsin; one missed cleavage allowed; fixed carbamidomethyl modification of cysteine, variable oxidation of methionine. Even though this particular method has its limitation as a sialic acid is required for the capturing, this could also be seen as an advantage in terms of mass spectrometry fragmentation patterns which are simplified due to the removal of the terminal sialic acid residue(s). For complex biological samples, this covalent capturing method with subsequent MS analysis is advantageous as the vast majority of non-glycosylated peptides is removed.[96] For further references on the analysis of tryptic O-glycopeptides see Table 1.

| Mass spectrometric approach | Sample | Achievements | Reference |
|---|---|---|---|
| *CID fragmentation* | | | |
| ESI-Q-TOF-$MS^2$ | Synthetic MUC2 glycopeptides | Identification of glycan attachment sites | 62,101 |
| FTICR-MS SORI-CID-$MS^2$ and $MS^3$ | Urinary O-glycopeptides from Schindler's disease | Negative-mode MS/MS for glycan structural elucidation | 119,120 |
| LC-ESI-IT-CID-$MS^2$; MALDI-TOF/TOF-$MS^2$ | Apolipoprotein CIII glycopeptides | Free O-glycopeptides with aberrant glycosylation found in urine of schistosomiasis patients | 121 |
| ESI-IT-CID-$MS^2$ | IgA1 hinge region tryptic O-glycopeptides | Glycoprofiling, data analysis using the GlycoSpectrumScan online tool | 122 |
| LC-ESI-IT-CID-$MS^2$ | Human IgG1 monoclonal antibody and CTLA4-Ig fusion protein | Identification of a novel O-glycosylation site on tryptic peptides from CTLA4-Ig fusion protein | 123 |
| nano-ESI-QQQ-MS and QTOF-MS / CID-$MS^2$ | Glycopeptides from Recombinant F-spondin | Identification of novel O-fucosylation sites | 124 |
| Nano-ESI-QTOF-MS / CID $MS^2$ | synthetic bisglycosylated MUC1 peptide 25-mer | Comparison of fragmentation spectra on nano-ESI-QTOF-MS/MS and PSD- MALDI-TOF-MS | 98 |
| ESI-QTOF-CID-$MS^2$ | Tryptic digests with subdigestions of Thrombospondin-1 | Identification of three site that carry the O-linked disaccharide Glc-Fuc-O Ser/Thr | 125 |
| Chip based nano-ESI-QTOF-MS / CID-$MS^2$ | Pronase E glycopeptides from bovine RNAse B, bovine Lactoferrin, bovine κ-casein, human IgG | Non-specifically cleaved glycopeptides can be analysed in a quantitative, isomer-sensitive and site-specific manner | 126 |

**Table 1.** Mass spectrometric analysis of O-glycosylation at the glycopeptides level.

| Mass spectrometric approach | Sample | Achievements | Reference |
|---|---|---|---|
| nano-ESI-QTOF-MS / CID-$MS^2$ | Peptides originating from the thrombospondin-1 repeat | Protocol optimized for the identification of O-fucosylation sites | 127 |
| MALDI-TOF-MS; ESI-MS / CID-$MS^2$ | O-glycosylated synthetic mucin-derived peptides | Peptide sequencing by CID-MS/MS enables the determination of the formerly O-glycosylated sites | 81 |
| ESI-QTOF-$MS^2$ | Glycopeptides and Clostripain-cleaved MUC1 glycopeptides from human breast cancer cell lines | Breast cancer cell line seems to glycosylate the MUC1 tandem repeat peptide at higher density than lactating breast epithelia | 128 |
| nanoESI-QTOF-MS/ CID-$MS^2$ | Pronase E derived glycopeptides from a mixture of bovine lactoferrin, bovine kappa casein and bovine fetuin | Protocol for separating and simultaneously characterizing glycopeptides from a mixture | 129 |
| UPLC-ESI-QTOF MS/ CID-$MS^2$ | Urine from individuals with urinary tract infections and healthy controls | Identified and structural characterization of a unique C-terminal O-glycopeptide of the human fibrinogen alpha-chain | 130 |
| nanoLC-LTQ-MS/ CID-$MS^2$ and -$MS^3$ | Tryptic α-dystroglycan glycopeptides | Developed a workflow and identified 91 glycopeptides that allowed us to assign 16 specific O-glycosylated residues | 131 |
| MALDI-TOF-MS and MALDI low-energy CID $MS^2$ on a QIT-TOF-MS | O-glycopeptides from recombinant human erythropoietin (rhEPO) from various suppliers obtained by digestion with a cleavage enzyme mix | Comparison of the O-glycopeptides revealed significant differences of the glycosylation patterns between the different rhEPO samples | 102 |

**Table 1.** (Continued).

| Mass spectrometric approach | Sample | Achievements | Reference |
|---|---|---|---|
| ZIC-HILIC-ESI-LIT-TOF-MS/ CID-$MS^2$ | Glu-C digested rhEPO | ZIC-HILIC-MS method for the characterization of N- as well as O-glycosylation | 132 |
| LC-LTQ MS / CID-$MS^2$ | Tryptic digests of muscle glycoproteins | Mucin-like domain of DGFc4 bared these modifications: the phosphorylated trisaccharide in conjunction with Hex-HexNAc-Hex, HexNAc-Hex, or Hex | 133 |
| nanoLC-ESI-MS/ CID-MS2 and -$MS^3$ | Proteinase K generated glycopeptides of (asialo) fetuin | Protocol for HILIC separation and MS characterization of N- as well as O-glycopeptides | 94 |
| *ETD/ ECD fragmentation* | | | |
| nanoLC-ESI-LIT-FTICR-MS/ HCD- and ETD-$MS^2$ | Three different human cell lines | Structural elucidation of the O-glycome within the cell lines | 118 |
| nanoLC-ESI-FTICR-$MS^2$ and -$MS^3$; combination of external accumulation (XA) with ECD | A 60-residue human-derived peptide and some synthetic peptides | XA-ECD allows shortened data acquisition time and facilitates the analysis of labile molecules | 103 |
| FTICR-MS / ECD-$MS^2$ | O-glycosylated synthetic mucin-derived peptides | "Mild" ECD protocol suitable for analysis of such labile species as glycosylated polypeptides | 104 |
| FTICR-MS / ECD-$MS^2$ | Standard peptides as well as synthetic glycosylated species | Improved ECD protocol for the analysis and characterization of O-glycosylated peptides | 105 |
| nano-ESI-FTICR-MS / ECD-$MS^2$ | O-fucosylated glycopeptides from thrombospondin-1 and properdin | Protocol for characterization of O-fucosylated peptides | 109 |

**Table 1.** (Continued).

| Mass spectrometric approach | Sample | Achievements | Reference |
|---|---|---|---|
| ESI-Q-FTICR-MS / ECD-MS2 | Synthetic IgA1 HR peptide and glycopeptide variants and IgA1 HR from a naturally Gal-deficient IgA1 myeloma protein | Successful identification of sites of attachment of O-glycans in the HR of a IgA1 myeloma protein | 110 |
| ESI-Q-FTICR-MS / activated ion ECD $MS^2$ | IgA1 myeloma proteins from patient sera | Protocol for characterization of site specific glycosylation using only with small amounts of total protein | 111 |
| MALDI-TOF-MS and ESI-QTOF-MS / CID-$MS^2$ | Tryptic (glyco)peptides derived from hepatitis B virus | Characterization of the overall glycosylation of all the subdomains within the virus | 100 |
| nanoLC- LTQ-FTICR-MS / ECD-$MS^2$ | Clostripain derived glycopeptides from MUC1-Ig fusion protein | Determination of both the average O-glycosylation site occupancy and the occupancy of individual glycosylation sites in the TR of MUC1 glycoproteins | 134 |
| LC-ESI-FTICR-MS; ECD-$MS^2$ and ETD-$MS^2$ | IgA-specific protease and trypsic digest of polymeric form of IgA1 myeloma protein from patient plasma | A new clinically relevant approach that requires ECD/electron transfer dissociation-type fragmentation to analyze clustered sites of O-glycosylation | 115 |
| Direct infusion on ESI-IT-MS/ ETD-$MS^2$ | Tryptic peptide of the Fab region of human IgG | Identification of O-fucose on the antibody | 135 |
| MALDI-TOF-MS and ESI-IT-MS / ETD-$MS^2$ | Enzymatic digestion by trypsin and lysylendopeptidase of IgA and hemopexin from sera of RA patients and healthy controls | Identification of an abnormality in the O-glycosylation of IgA associated with rheumatoid arthritis | 136 |

**Table 1.** (Continued).

| Mass spectrometric approach | Sample | Achievements | Reference |
|---|---|---|---|
| *CID and ETD/ECD fragmentation* | | | |
| nanoLC-MS / CID- and ETD-$MS^2$ | Secreted APP695 produced in CHO cells | Identification and structural elucidation of three new O-glycosylation sites on APP695 | 93 |
| ESI-IT-MS / CID- and ETD-$MS^2$ | Synthetic MUC1 glycopeptides carrying different degree of O-glycosylation | Site specific characterisation of mucin-type glycopeptides containing highly densed O-glycan clusters | 113 |
| MALDI-TOF-MS and ESI-TOF-MS / ETD- and CID-$MS^2$ | Cellobiohydrolase I&II and Endoglucanase I&II from *Trichoderma reesei* | Combining info from different analytical approaches allow to elucidation of heavily glycosylated O-linked glycopeptides | 114 |
| ESI-LIT-TOF-MS / tandem LITs for CID- and ECD-$MS^2$ and $MS^3$ | Synthetic O-glycopeptides | Negative-ion CID and positive-ion ECD MSn spectra play complementary roles in the direct structural analysis of O-glycopeptides | 97 |
| nanoLC-LTQ-Orbitrap-MS / ETD- and CID-$MS^2$ | O-glycopeptides enriched from bovine serum | Advanced protocols combining ETD and CID fragmentation spectra using Protein Prospector for identification | 92,95,117 |
| nano-LC- ESI-QTOF-MS / CID- and ETD-$MS^2$ | Tryptic glycopeptides of AP 180 | Tryptic glycopeptides of AP 180 | 137 |
| nanoLC-ESI-FTICR-MS / CID- and ECD-$MS^2$ and $MS^3$ | Peptides and glycopeptides from amyloid precursor protein/ amyloid β-peptides in CSF | Identification of novel O-glycosylation sites in threonins and tyrosine 10. Tyr glycosylation might be involved in AD progression | 108 |

**Table 1.** (Continued).

| Mass spectrometric approach | Sample | Achievements | Reference |
|---|---|---|---|
| nanoLC-ESI-LTQ-FTICR-MS / CID- and ECD-$MS^2$ and $MS^3$ | Tryptic glycopeptides from human urine | 63 O-linked glycopeptides from 53 glycoproteins were characterized of which 29 novel sites | 138 |
| FTICR-MS/ CID- and ECD-$MS^2$ | Synthetic mucin derived O-glycopeptides | Protocol for the structural elucidation of mucin derived O-glycopeptides by ECD | 5 |
| nanoLC- LIT-ETD- and CID-$MS^2$ | Synthetic glycopeptides and proteins of liver in response to traumahemorrhage and resuscitation in a rat model | Protocol using high affinity pan-specific antibodies for exploring the O-GlcNAc proteome | 139 |
| LC-ESI-FTICR-MS / CID- and ECD-MS2 and $MS^3$ | Human CSF samples | Qualitative information of N- and O-linked glycosylation core structures and their attachment sites in sialylated glycoproteins in a proteomic mixture | 96 |
| *Other MS techniques/ methods* | | | |
| MALDI-TOF-PSD-$MS^2$ | Synthetic MUC4 glycopeptides | For some glycopeptides observation of glycosylated fragment ions allowing site assignment | 101 |
| MALDI-ISD-MS | Synthetic glycopeptides, bovine asialofetuin and k-caseine | ISD MALDI MS is suitable for site determination of complex O-glycans at the protein level | 140 |
| Linear MALDI-TOF-MS and MALDI-TOF-PSD-$MS^2$ | Clostripain-generated glycopeptides from isolated human milk mucin | Characterization of major glycans from human skim milk MUC1 | 141 |

**Table 1.** (Continued).

| Mass spectrometric approach | Sample | Achievements | Reference |
|---|---|---|---|
| nano-ESI-Q-FTICR-MS2/ positive- and negative-mode IRMPD-MS$^2$ | Pronase E-derived glycopeptides from bovine fetuin A and bovine kappa-casein | Protocol for the IRMPD of the deprotonated form of O-glycopeptides | 142 |
| CE-ESI-FTICR-MS | recombinant human chorionic gonadotrophin | Protocol for high-resolution CE-LTQ-FT-MS for the profiling of the intact glycoforms of r-RhCG | 143 |

**Table 1.** (Continued).

## Glycopeptides generated by use of unspecific proteinases

Lately, the use of unspecific proteinases such as Pronase E (proteinase mixture from Streptomyces griseus) or Proteinase K (from Tritirachium album) for glycoprotein cleavage has been shown to be useful for in-depth analysis of the site-specific N- and/or O-glycosylation of individual glycoproteins in many cases.[94,114,129,142,144-149]

Pronase E was immobilized on Sepharose beads for digesting two model proteins (bovine fetuin A and bovine kappa-casein) prior to nanoLC-FTICR-MS applying the infrared multiphoton dissociation (IRMPD) mode, which allows an assortment of O-linked glycopeptide ions encompassing various compositions and charge states.[142] In particular, IRMPD of the deprotonated form of O-glycopeptides has been shown to produce informative side chain cleavages from unoccupied serine and threonine residues, allowing determination of whether a serine or threonine residue is glycosylated. A combination of positive mode and negative mode MS2 was found to provide conclusive assignment of O-glycosites.[142]

In another example, Christiansen et al.[114] showed in their paper that ETD mass spectrometry of the protease-digested and HILIC-enriched highly glycosylated

domains of different glycoproteins allows determining the overall oligosaccharide heterogeneity.[114] Different enzymes have been used for digestion such as trypsin, AspN but also some unspecific proteinases such as proteinase K and the so-called PreTAQ. The main aim of this particular work was to develop complementary strategies to obtain structural information on the glycosylation of these difficult to analyze heterogeneously O-glycosylated domains. However, the O-glycan site occupancy of the mucin-like domains, with the modification intact, requires the determination of the peptide sequence. Clearly, ETD MS/MS has the potential to assign these O-glycosylation sites and at the same time give information on the peptide sequence too. It was shown that by using different types of analytical approaches and putting the information together, many of the structural details of the glycan heterogeneity can be determined on these heavily glycosylated O-linked glycopeptides.[114]

## Analysis of synthetic glycopeptides

A more technical approach was undertaken by Thaysen-Anderson et al.[113] who used synthetic glycopeptides from the human mucin-1 (MUC-1) to test the use of ETD for the site-specific characterization of densely glycosylated mucin type O-glycopeptides. High peptide sequence coverage and confident glycosylation site determination were obtained from the infusion of individual glycopeptides, and from glycopeptide mixtures separated using two different online LC approaches (RP and porous graphitized carbon, PGC), using ETD-IT-MS2. Interestingly, highly sialylated glycopeptides showed poorer ETD fragmentation, which limits the applicability of the approach as mucin type O-glycans have a naturally high abundance of sialic acid residues. On the other hand, the capability to characterise the glycosylation sites of

mixtures of densely O-glycosylated mucin-type peptides while keeping the attached glycans structures intact during ETD fragmentation is amenable. Lately, synthetic O-glycopeptides served in some cases as excellent model systems to optimize various MS approaches (see Table 1 for details).

## Naturally occurring O-glycopeptides

Recently, it was found that O-glycopeptides can be used to distinguish between healthy and diseased state: for example Balog et al.[121] identified a group of novel aberrantly O-glycosylated peptide derived from the C-terminus of human apolipoprotein C-III (apoC-III) in the urine of *S. mansoni*-infected individuals which was absent for healthy controls. The O-glycan portions identified appeared to represent highly fucosylated core 2 structures.

In another comparative urinary metabolomic study of patients suffering from urinary tract infection (UTI), Pacchiarotta et al.[130] identified a fibrinogen α-chain O-glycopeptide decorated with a di-sialylated T-antigen as the major discriminator between UTI patients and controls. Both studies give new insight into the (glyco)biology that seems to play a role in the diseased state. The identification of such O-glycosylation based discriminators indicates the potential of urinary glycopeptides as biomarkers for infections.

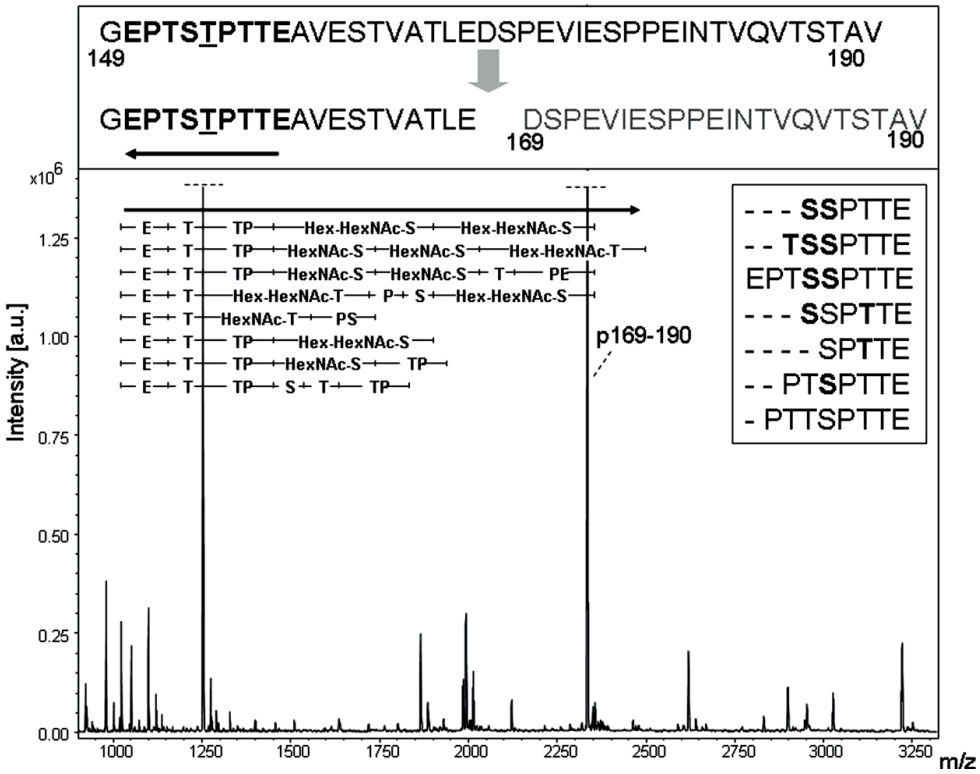



**Figure 6.** Top-down glycoprotein sequencing of bovine k-caseine. The sample was applied in sDHB matrix and analyzed by ISD-MALDI-MS. Sequences shown on top of the figure refer to the C-terminal domain of the protein, which spontaneously fragments in the ion source to yield p169-190 and a correspondingly truncated k-caseinoglycopeptide. The arrows indicate the reading direction of the amino acid sequence of the z-ion series annotated in the mass spectrum. The insert shows the readable section of clustered Ser/Thr (p150-158). The gray-shaded serine indicates a sequence ambiguity with published data. Glycosylated Ser/Thr residues are highlighted in bold. Taken from Hanisch et al.[140] with permission.

Thirty seven human amyloid precursor protein O-glycopeptides derived from CSF (attached to Thr(−39, −21, −20, and −13), with sialylated core 1 like O-glycans) and, unexpectedly, a series of 27 glycopeptides (the Aβ1-X series, where X was 20 (DAEFRHDSGYEVHHQKLVFF), which were all uniquely glycosylated on Tyr10 by ((Neu5Ac)1-2Hex-(Neu5Ac)-HexNAc-O- structures), were recently identified (Halim et al., 2011). An increase of up to 2.5 times of Tyr10 glycosylated peptides in CSF in

six Alzheimer's disease (AD) patients compared to seven non-AD patients was observed. These sialylated O-glycans, including that of a Tyr residue, the first in a mammalian protein, may modulate human amyloid precursor protein processing, and therefore influence the amyloidogenic pathway associated with AD.

The expected proteolytic resistance of the O-glycosylated peptide stretches may lead to the accumulation of these glycopeptide moieties in degradation pathways, in particular in pathological situations. Hopefully, with the advances in the field of O-glycoproteomics many more O-glycopeptides can be identified within the near future that may be important in (monitoring) disease.[150]

## Perspectives

One of the major challenges in the analysis of O-glycans is the unwanted side reaction occurring once the oligosaccharides are released from the protein backbone (named 'peeling'[69]). Hopefully, there will be ways indentified to reduce/abolish this 'peeling' reaction and/or combine the release of O-glycans with a suitable labelling technique. Some recent papers have described first steps towards this goal.[55,67,68,71,151] The efficient blocking of the reducing end of the formerly O-linked oligosaccharide appears to be an important aspect in improving the stability of the O-glycans.

Rather novel techniques for O-glycopeptide analysis such as the radical-mediated mass spectrometric fragmentation methods ETD and ECD are becoming more and more popular for site specific glycoprotein analysis and have expanded the analytical options in glycoproteomics tremendously (see the ECD/ETD Fragmentation of Glycopeptides section). These methods (in particular ETD) will,

probably in combination with CID fragmentation, allow a very detailed analysis of more glycoproteins in the future.

Other important aspects in the O-glycoproteomics field are spectral assignment and data analysis. In this respect, a very interesting collaboration of researchers is the UniCarbDB platform (http://www.unicarb-db.com/) which is currently focused on the mass spectrometric data and structural assignment based on fragmentation data. The aim is to generate a assigned database of assigned spectra and structures supporting structural elucidation.[152]

Recent publications show that combinatorial approaches or glyco-engineering can help to get a picture of the O-glycoproteome of the sample of interest.[96,118] While these approaches are successful for simple or simplified O-glycosylation, they might not be suitable to deal with peptides carrying large and structurally diverse O-glycans, requiring manual interpretation of the corresponding fragmentation spectra.

One very elegant approach has recently been published, in which top-down sequencing on the protein level by MALDI-MS based on the in-source decay (ISD) of intact glycoproteins was induced by hydrogen radical transfer from the matrix.[140] It allows a ladder sequencing from both termini with assignment of O-glycosylation sites based on intense c-, y- and z-type ions. In Figure 6 the top-down glycoprotein sequencing of bovine k-caseine is shown as an example. The feasibility of ISD-MALDI-MS in the localization of O-glycosylation sites was clearly demonstrated for several proteins as fragments are formed under conditions that leave O-glycosidic bonds largely unaffected. This method gives comparable information to that obtained with post source decay (PSD) MALDI or ETD-ESI-MS, which shows its potential for the site determination of complex O-glycans at the protein level. After further

optimization of this approach, it is expected to present an attractive and powerful tool for the site determination of complex O-glycans at the protein level.[140]

The above mentioned technical approaches are already major improvements/advances to study O-glycosylation in greater detail. We believe that also in O-glycoproteomics we have only seen the tip of the glycobiology iceberg,[153] and that there will be major and significant discoveries ahead of us due to the combination/ further improvements of all the latest technical advances.

## Reference List:

(1) Spiro, R. G. *Glycobiology* **2002**, *12*, 43R-56R.

(2) Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. *Essential of Glycobiology*, 2 ed.; Cold Spring Harbor Laboratory Press, 2009.

(3) Gill, D. J.; Clausen, H.; Bard, F. *Trends Cell Biol* **2011**, *21*, 149-158.

(4) Ten Hagen, K. G.; Fritz, T. A.; Tabak, L. A. *Glycobiology* **2003**, *13*, 1R-16R.

(5) Mormann, M.; Paulsen, H.; Peter-Katalinic, J. *Eur. J. Mass Spectrom. (Chichester, Eng)* **2005**, *11*, 497-511.

(6) Park, J. H.; Nishidate, T.; Kijima, K.; Ohashi, T.; Takegawa, K.; Fujikane, T.; Hirata, K.; Nakamura, Y.; Katagiri, T. *Cancer Res* **2010**, *70*, 2759-2769.

(7) Taylor-Papadimitriou, J.; Burchell, J.; Miles, D. W.; Daziel, M. *Biochimica et Biophysica Acta* **1999**, *1455*, 301-313.

(8) Pedersen, J. W.; Blixt, O.; Bennett, E. P.; Tarp, M. A.; Dar, I.; Mandel, U.; Poulsen, S. S.; Pedersen, A. E.; Rasmussen, S.; Jess, P.; Clausen, H.; Wandall, H. H. *Int. J. Cancer* **2011**, *128*, 1860-1871.

(9) Stalnaker, S. H.; Stuart, R.; Wells, L. *Curr. Opin. Struct. Biol* **2011**, *21*, 603-609.

(10) Haltiwanger, R. S.; Lowe, J. B. *Annu. Rev. Biochem* **2004**, *73*, 491-537.

(11) Sethi, M. K.; Buettner, F. F.; Ashikov, A.; Krylov, V. B.; Takeuchi, H.; Nifantiev, N. E.; Haltiwanger, R. S.; Gerardy-Schahn, R.; Bakker, H. *J. Biol. Chem* **2012**, *287*, 2739-2748.

(12) Lazarus, B. D.; Love, D. C.; Hanover, J. A. *Int. J. Biochem. Cell Biol* **2009**, *41*, 2134-2146.

(13) Geyer, H.; Geyer, R. *Biochimica et Biophysica Acta* **2006**, *1764*, 1853-1869.

(14) Harvey, D. J. *Expert. Rev. Proteomics* **2005**, *2*, 87-101.

(15) Zaia, J. *OMICS* **2010**, *14*, 401-418.

(16) Hulsmeier, A. J.; Welti, M.; Hennet, T. *Methods Enzymol* **2011**, *491*, 163-182.

(17) Wada, Y.; Dell, A.; Haslam, S. M.; Tissot, B.; Canis, K.; Azadi, P.; Backstrom, M.; Costello, C. E.; Hansson, G. C.; Hiki, Y.; Ishihara, M.; Ito, H.; Kakehi, K.; Karlsson, N.; Hayes, C. E.; Kato, K.; Kawasaki, N.; Khoo, K. H.; Kobayashi, K.; Kolarich, D.; Kondo, A.; Lebrilla, C.; Nakano, M.; Narimatsu, H.; Novak, J.; Novotny,

M. V.; Ohno, E.; Packer, N. H.; Palaima, E.; Renfrow, M. B.; Tajiri, M.; Thomsson, K. A.; Yagi, H.; Yu, S. Y.; Taniguchi, N. *Mol. Cell Proteomics* **2010**, *9*, 719-727.
(18) Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M. *Anal. Bioanal. Chem* **2010**, *397*, 3457-3481.
(19) North, S. J.; Hitchen, P. G.; Haslam, S. M.; Dell, A. *Curr. Opin. Struct. Biol* **2009**, *19*, 498-506.
(20) Peter-Katalinic, J. *Methods Enzymol* **2005**, *405*, 139-171.
(21) Dwek, R. A.; Edge, C. J.; Harvey, D. J.; Wormald, M. R.; Parekh, R. B. *Annu. Rev. Biochem* **1993**, *62*, 65-100.
(22) Carlson, D. M. *J. Biol. Chem* **1968**, *243*, 616-626.
(23) Iyer, R. N.; Carlson, D. M. *Arch. Biochem. Biophys* **1971**, *142*, 101-105.
(24) Cooper, C. A.; Packer, N. H.; Redmond, J. W. *Glycoconj. J* **1994**, *11*, 163-167.
(25) Brooks, R. D. *A kinetic study of the Rate of Cleavage of the Glycosidic Bond of Methyl-beta-Glucopyranoside in an Alkaline Medium*. The Institute of Paper Chemistry, Appleton, Wisconsin.1966.
(26) Corbett, W. M.; Kenner, J. *Journal of the Chemical Society* **1955**, 1431-1435.
(27) Greville, M.; Richards, G. N. *Journal of the Chemical Society* **1960**, 1924-1931.
(28) MacLaurin, D. J. *A study of some reaction rates in the homogeneous system water-sodium hydroxide-cellobiose*. The Institute of Paper Chemistry, Appleton, Wisconsin.1969.
(29) Richards, G. N.; Sephton, H. H. *Journal of the Chemical Society* **1957**, 4492-4499.
(30) WHISTLER, R. L.; BEMILLER, J. N. *Adv. Carbohydr. Chem* **1958**, *13*, 289-329.
(31) Kotani, N.; Takasaki, S. *Anal. Biochem* **1997**, *252*, 40-47.
(32) Lloyd, K. O.; Burchell, J.; Kudryashov, V.; Yin, B. W.; Taylor-Papadimitriou, J. *J. Biol. Chem* **1996**, *271*, 33325-33334.
(33) Doohan, R. A.; Hayes, C. A.; Harhen, B.; Karlsson, N. G. *J. Am. Soc. Mass Spectrom* **2011**, *22*, 1052-1062.
(34) Faid, V.; Chirat, F.; Seta, N.; Foulquier, F.; Morelle, W. *Proteomics* **2007**, *7*, 1800-1813.
(35) Ismail, M. N.; Stone, E. L.; Panico, M.; Lee, S. H.; Luu, Y.; Ramirez, K.; Ho, S. B.; Fukuda, M.; Marth, J. D.; Haslam, S. M.; Dell, A. *Glycobiology* **2011**, *21*, 82-98.
(36) Schulz, B. L.; Packer, N. H.; Karlsson, N. G. *Anal. Chem* **2002**, *74*, 6088-6097.

(37) Taylor, A. M.; Holst, O.; Thomas-Oates, J. *Proteomics* **2006**, *6*, 2936-2946.
(38) Cooke, C. L.; An, H. J.; Kim, J.; Solnick, J. V.; Lebrilla, C. B. *Anal. Chem* **2007**, *79*, 8090-8097.
(39) Robijn, M. L.; Koeleman, C. A.; Wuhrer, M.; Royle, L.; Geyer, R.; Dwek, R. A.; Rudd, P. M.; Deelder, A. M.; Hokke, C. H. *Mol. Biochem. Parasitol* **2007**, *151*, 148-161.
(40) Kawasaki, N.; Ohta, M.; Hyuga, S.; Hyuga, M.; Hayakawa, T. *Anal. Biochem* **2000**, *285*, 82-91.
(41) Thomsson, K. A.; Prakobphol, A.; Leffler, H.; Reddy, M. S.; Levine, M. J.; Fisher, S. J.; Hansson, G. C. *Glycobiology* **2002**, *12*, 1-14.
(42) Karlsson, N. G.; Schulz, B. L.; Packer, N. H. *J. Am. Soc. Mass Spectrom* **2004**, *15*, 659-672.
(43) Larsson, J. M.; Karlsson, H.; Sjovall, H.; Hansson, G. C. *Glycobiology* **2009**, *19*, 756-766.
(44) Ciucanu, I.; Costello, C. E. *J. Am. Chem. Soc* **2003**, *125*, 16213-16219.
(45) Jang-Lee, J.; North, S. J.; Sutton-Smith, M.; Goldberg, D.; Panico, M.; Morris, H.; Haslam, S.; Dell, A. *Methods Enzymol* **2006**, *415*, 59-86.
(46) Ciucanu, I.; Kerek, F. *Carbohydrate Research* **1984**, *131*, 209-217.
(47) Geyer, R.; Geyer, H. *Methods Enzymol* **1994**, *230*, 86-108.
(48) Morelle, W.; Michalski, J. C. *Nat. Protoc* **2007**, *2*, 1585-1602.
(49) Morelle, W.; Faid, V.; Chirat, F.; Michalski, J. C. *Methods Mol. Biol* **2009**, *534*, 5-21.
(50) Robbe, C.; Michalski, J. C.; Capon, C. *Methods Mol. Biol* **2006**, *347*, 109-123.
(51) Hanisch, F. G.; Uhlenbruck, G.; Peter-Katalinic, J.; Egge, H.; Dabrowski, J.; Dabrowski, U. *J. Biol. Chem* **1989**, *264*, 872-883.
(52) Egge, H.; Peter-Katalinic, J. *Mass Spectrometry Reviews* **1987**, *6*, 331-393.
(53) Bleckmann, C.; Geyer, H.; Lieberoth, A.; Splittstoesser, F.; Liu, Y.; Feizi, T.; Schachner, M.; Kleene, R.; Reinhold, V.; Geyer, R. *Biol. Chem* **2009**, *390*, 627-645.
(54) Parry, S.; Wong, N. K.; Easton, R. L.; Panico, M.; Haslam, S. M.; Morris, H. R.; Anderson, P.; Klotz, K. L.; Herr, J. C.; Diekman, A. B.; Dell, A. *Glycobiology* **2007**, *17*, 1120-1126.
(55) Maniatis, S.; Zhou, H.; Reinhold, V. *Anal. Chem* **2010**, *82*, 2421-2425.
(56) Hu, P.; Shimoji, S.; Hart, G. W. *FEBS Lett* **2010**, *584*, 2526-2538.

(57) Chai, W.; Feizi, T.; Yuen, C. T.; Lawson, A. M. *Glycobiology* **1997**, *7*, 861-872.
(58) Yamada, K.; Hyodo, S.; Kinoshita, M.; Hayakawa, T.; Kakehi, K. *Anal. Chem* **2010**, *82*, 7436-7443.
(59) Huang, Y.; Mechref, Y.; Novotny, M. V. *Anal. Chem* **2001**, *73*, 6063-6069.
(60) Yu, G.; Zhang, Y.; Zhang, Z.; Song, L.; Wang, P.; Chai, W. *Anal. Chem* **2010**, *82*, 9534-9542.
(61) Miura, Y.; Kato, K.; Takegawa, Y.; Kurogochi, M.; Furukawa, J.; Shinohara, Y.; Nagahori, N.; Amano, M.; Hinou, H.; Nishimura, S. *Anal. Chem* **2010**, *82*, 10021-10029.
(62) Domon, B.; Costello, C. E. *Glycoconjugate journal* **1988**, *5*, 397-409.
(63) Goetz, J. A.; Novotny, M. V.; Mechref, Y. *Anal. Chem* **2009**, *81*, 9546-9552.
(64) Kang, P.; Mechref, Y.; Klouckova, I.; Novotny, M. V. *Rapid Commun. Mass Spectrom* **2005**, *19*, 3421-3428.
(65) Kang, P.; Mechref, Y.; Novotny, M. V. *Rapid Commun. Mass Spectrom* **2008**, *22*, 721-734.
(66) Zauner, G.; Koeleman, C. A.; Deelder, A. M.; Wuhrer, M. *Biochim. Biophys. Acta* **2012**, *1820*, 1420-1428.
(67) Wang, C.; Fan, W.; Zhang, P.; Wang, Z.; Huang, L. *Proteomics* **2011**, *11*, 4229-4242.
(68) Furukawa, J.; Fujitani, N.; Araki, K.; Takegawa, Y.; Kodama, K.; Shinohara, Y. *Anal. Chem* **2011**, *83*, 9060-9067.
(69) Merry, A. H.; Neville, D. C.; Royle, L.; Matthews, B.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M. *Anal. Biochem* **2002**, *304*, 91-99.
(70) Patel, T.; Bruce, J.; Merry, A.; Bigge, C.; Wormald, M.; Jaques, A.; Parekh, R. *Biochemistry* **1993**, *32*, 679-693.
(71) Kozak, R. P.; Royle, L.; Gardner, R. A.; Fernandes, D. L.; Wuhrer, M. *Anal. Biochem* **2012**, *423*, 119-128.
(72) Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A.; Haslam, S. M. *J. Proteome. Res* **2008**, *7*, 1650-1659.
(73) Maass, K.; Ranzinger, R.; Geyer, H.; von der Lieth, C. W.; Geyer, R. *Proteomics* **2007**, *7*, 4435-4444.
(74) Campbell, M. P.; Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. *Bioinformatics* **2008**, *24*, 1214-1216.

(75) Jansson, P. E.; Stenutz, R.; Widmalm, G. *Carbohydr. Res* **2006**, *341*, 1003-1010.
(76) Lutteke, T.; Bohne-Lang, A.; Loss, A.; Goetz, T.; Frank, M.; von der Lieth, C. W. *Glycobiology* **2006**, *16*, 71R-81R.
(77) Jensen, P. H.; Kolarich, D.; Packer, N. H. *FEBS J* **2010**, *277*, 81-94.
(78) Rademaker, G. J.; Haverkamp, J.; Thomas-Oates, J. *Organic Mass Spectrometry* **1993**, *28*, 1536-1541.
(79) Rademaker, G. J.; Pergantis, S. A.; Blok-Tip, L.; Langridge, J. I.; Kleen, A.; Thomas-Oates, J. E. *Anal. Biochem* **1998**, *257*, 149-160.
(80) Hanisch, F. G.; Jovanovic, M.; Peter-Katalinic, J. *Analytical biochemistry* **2001**, *290*, 47-59.
(81) Mirgorodskaya, E.; Hassan, H.; Clausen, H.; Roepstorff, P. *Analytical chemistry* **2001**, *73*, 1263-1269.
(82) Zheng, Y.; Guo, Z.; Cai, Z. *Talanta* **2009**, *78*, 358-363.
(83) Wells, L.; Vosseller, K.; Cole, R. N.; Cronshaw, J. M.; Matunis, M. J.; Hart, G. W. *Molecular & cellular proteomics : MCP* **2002**, *1*, 791-804.
(84) Vosseller, K.; Hansen, K. C.; Chalkley, R. J.; Trinidad, J. C.; Wells, L.; Hart, G. W.; Burlingame, A. L. *Proteomics* **2005**, *5*, 388-398.
(85) Czeszak, X.; Ricart, G.; Tetaert, D.; Michalski, J. C.; Lemoine, J. *Rapid communications in mass spectrometry : RCM* **2002**, *16*, 27-34.
(86) Hansen, J. E.; Lund, O.; Tolstrup, N.; Gooley, A. A.; Williams, K. L.; Brunak, S. *Glycoconjugate journal* **1998**, *15*, 115-130.
(87) Hansen, J. E.; Lund, O.; Nielsen, J. O.; Hansen, J. E.; Brunak, S. *Nucleic acids research* **1996**, *24*, 248-252.
(88) Chen, Y. Z.; Tang, Y. R.; Sheng, Z. Y.; Zhang, Z. *BMC bioinformatics* **2008**, *9*, 101.
(89) Medzihradszky, K. F.; Gillece-Castro, B. L.; Settineri, C. A.; Townsend, R. R.; Masiarz, F. R.; Burlingame, A. L. *Biomedical & environmental mass spectrometry* **1990**, *19*, 777-781.
(90) Medzihradszkyaff, K. F.; Gillece-Castroaff, B. L.; Townsendaff, R. R.; Burlingameaff, A. L.; Hardyaff, M. R. *Journal of the American Society for Mass Spectrometry* **1996**, *7*, 319-328.

(91) Settineri, C. A.; Medzihradszky, K. F.; Masiarz, F. R.; Burlingame, A. L.; Chu, C.; George-Nascimento, C. *Biomedical & environmental mass spectrometry* **1990**, *19*, 665-676.
(92) Darula, Z.; Chalkley, R. J.; Baker, P.; Burlingame, A. L.; Medzihradszky, K. F. *European journal of mass spectrometry (Chichester, England)* **2010**, *16*, 421-428.
(93) Perdivara, I.; Petrovich, R.; Allinquant, B.; Deterding, L. J.; Tomer, K. B.; Przybylski, M. *Journal of proteome research* **2009**, *8*, 631-642.
(94) Zauner, G.; Koeleman, C. A.; Deelder, A. M.; Wuhrer, M. *Journal of separation science* **2010**, *33*, 903-910.
(95) Darula, Z.; Chalkley, R. J.; Lynn, A.; Baker, P. R.; Medzihradszky, K. F. *Amino acids* **2011**, *41*, 321-328.
(96) Nilsson, J.; Ruetschi, U.; Halim, A.; Hesse, C.; Carlsohn, E.; Brinkmalm, G.; Larson, G. *Nature methods* **2009**, *6*, 809-811.
(97) Deguchi, K.; Ito, H.; Baba, T.; Hirabayashi, A.; Nakagawa, H.; Fumoto, M.; Hinou, H.; Nishimura, S. *Rapid communications in mass spectrometry : RCM* **2007**, *21*, 691-698.
(98) Hanisch, F. G.; Green, B. N.; Bateman, R.; Peter-Katalinic, J. *Journal of mass spectrometry : JMS* **1998**, *33*, 358-362.
(99) Wuhrer, M.; Catalina, M. I.; Deelder, A. M.; Hokke, C. H. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2007**, *849*, 115-128.
(100) Schmitt, S.; Glebe, D.; Alving, K.; Tolle, T. K.; Linder, M.; Geyer, H.; Linder, D.; Peter-Katalinic, J.; Gerlich, W. H.; Geyer, R. *The Journal of biological chemistry* **1999**, *274*, 11945-11957.
(101) Alving, K.; Paulsen, H.; Peter-Katalinic, J. *Journal of mass spectrometry : JMS* **1999**, *34*, 395-407.
(102) Stubiger, G.; Marchetti, M.; Nagano, M.; Grimm, R.; Gmeiner, G.; Reichel, C.; Allmaier, G. *Journal of separation science* **2005**, *28*, 1764-1778.
(103) Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnsen, A. H.; Zubarev, R. A. *Analytical chemistry* **2001**, *73*, 2998-3005.
(104) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Analytical chemistry* **1999**, *71*, 4431-4436.

(105) Mormann, M.; Peter-Katalinic, J. *Rapid communications in mass spectrometry : RCM* **2003**, *17*, 2208-2214.
(106) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Analytical chemistry* **2000**, *72*, 563-573.
(107) Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sorensen, E. S.; Zubarev, R. A. *Analytical chemistry* **2003**, *75*, 2355-2361.
(108) Halim, A.; Brinkmalm, G.; Ruetschi, U.; Westman-Brinkmalm, A.; Portelius, E.; Zetterberg, H.; Blennow, K.; Larson, G.; Nilsson, J. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 11848-11853.
(109) Mormann, M.; Macek, B.; Gonzalez de Peredo, A.; Hofsteenge, J.; Peter-Katalinic, J. *International Journal of Mass Spectrometry* **2004**, *234*, 11-21.
(110) Renfrow, M. B.; Cooper, H. J.; Tomana, M.; Kulhavy, R.; Hiki, Y.; Toma, K.; Emmett, M. R.; Mestecky, J.; Marshall, A. G.; Novak, J. *The Journal of biological chemistry* **2005**, *280*, 19136-19145.
(111) Renfrow, M. B.; Mackay, C. L.; Chalmers, M. J.; Julian, B. A.; Mestecky, J.; Kilian, M.; Poulsen, K.; Emmett, M. R.; Marshall, A. G.; Novak, J. *Analytical and bioanalytical chemistry* **2007**, *389*, 1397-1407.
(112) Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 9528-9533.
(113) Thaysen-Andersen, M.; Wilkinson, B. L.; Payne, R. J.; Packer, N. H. *Electrophoresis* **2011**, *32*, 3536-3545.
(114) Christiansen, M. N.; Kolarich, D.; Nevalainen, H.; Packer, N. H.; Jensen, P. H. *Analytical chemistry* **2010**, *82*, 3500-3509.
(115) Takahashi, K.; Wall, S. B.; Suzuki, H.; Smith, A. D. t.; Hall, S.; Poulsen, K.; Kilian, M.; Mobley, J. A.; Julian, B. A.; Mestecky, J.; Novak, J.; Renfrow, M. B. *Molecular & cellular proteomics : MCP* **2010**, *9*, 2545-2557.
(116) Halfinger, B.; Sarg, B.; Lindner, H. H. *Electrophoresis* **2011**, *32*, 3546-3553.
(117) Darula, Z.; Medzihradszky, K. F. *Molecular & cellular proteomics : MCP* **2009**, *8*, 2515-2526.

(118) Steentoft, C.; Vakhrushev, S. Y.; Vester-Christensen, M. B.; Schjoldager, K. T.; Kong, Y.; Bennett, E. P.; Mandel, U.; Wandall, H.; Levery, S. B.; Clausen, H. *Nature methods* **2011**, *8*, 977-982.
(119) Froesch, M.; Bindila, L.; Zamfir, A.; Peter-Katalinic, J. *Rapid communications in mass spectrometry : RCM* **2003**, *17*, 2822-2832.
(120) Froesch, M.; Bindila, L. M.; Baykut, G.; Allen, M.; Peter-Katalinic, J.; Zamfir, A. D. *Rapid communications in mass spectrometry : RCM* **2004**, *18*, 3084-3092.
(121) Balog, C. I.; Mayboroda, O. A.; Wuhrer, M.; Hokke, C. H.; Deelder, A. M.; Hensbergen, P. J. *Molecular & cellular proteomics : MCP* **2010**, *9*, 667-681.
(122) Deshpande, N.; Jensen, P. H.; Packer, N. H.; Kolarich, D. *Journal of proteome research* **2010**, *9*, 1063-1075.
(123) Bongers, J.; Devincentis, J.; Fu, J.; Huang, P.; Kirkley, D. H.; Leister, K.; Liu, P.; Ludwig, R.; Rumney, K.; Tao, L.; Wu, W.; Russell, R. J. *Journal of chromatography. A* **2011**, *1218*, 8140-8149.
(124) Gonzalez de Peredo, A.; Klein, D.; Macek, B.; Hess, D.; Peter-Katalinic, J.; Hofsteenge, J. *Molecular & cellular proteomics : MCP* **2002**, *1*, 11-18.
(125) Hofsteenge, J.; Huwiler, K. G.; Macek, B.; Hess, D.; Lawler, J.; Mosher, D. F.; Peter-Katalinic, J. *The Journal of biological chemistry* **2001**, *276*, 6485-6498.
(126) Hua, S.; Nwosu, C. C.; Strum, J. S.; Seipert, R. R.; An, H. J.; Zivkovic, A. M.; German, J. B.; Lebrilla, C. B. *Analytical and bioanalytical chemistry* **2012**, *403*, 1291-1302.
(127) Macek, B.; Hofsteenge, J.; Peter-Katalinic, J. *Rapid communications in mass spectrometry : RCM* **2001**, *15*, 771-777.
(128) Muller, S.; Alving, K.; Peter-Katalinic, J.; Zachara, N.; Gooley, A. A.; Hanisch, F. G. *The Journal of biological chemistry* **1999**, *274*, 18165-18172.
(129) Nwosu, C. C.; Seipert, R. R.; Strum, J. S.; Hua, S. S.; An, H. J.; Zivkovic, A. M.; German, B. J.; Lebrilla, C. B. *Journal of proteome research* **2011**, *10*, 2612-2624.
(130) Pacchiarotta, T.; Hensbergen, P. J.; Wuhrer, M.; van Nieuwkoop, C.; Nevedomskaya, E.; Derks, R. J.; Schoenmaker, B.; Koeleman, C. A.; van Dissel, J.; Deelder, A. M.; Mayboroda, O. A. *Journal of proteomics* **2012**, *75*, 1067-1073.
(131) Stalnaker, S. H.; Hashmi, S.; Lim, J. M.; Aoki, K.; Porterfield, M.; Gutierrez-Sanchez, G.; Wheeler, J.; Ervasti, J. M.; Bergmann, C.; Tiemeyer, M.; Wells, L. *The Journal of biological chemistry* **2010**, *285*, 24882-24891.

(132) Takegawa, Y.; Ito, H.; Keira, T.; Deguchi, K.; Nakagawa, H.; Nishimura, S. *Journal of separation science* **2008**, *31*, 1585-1593.
(133) Yoshida-Moriguchi, T.; Yu, L.; Stalnaker, S. H.; Davis, S.; Kunz, S.; Madson, M.; Oldstone, M. B.; Schachter, H.; Wells, L.; Campbell, K. P. *Science (New York, N.Y.)* **2010**, *327*, 88-92.
(134) Sihlbom, C.; van Dijk Hard, I.; Lidell, M. E.; Noll, T.; Hansson, G. C.; Backstrom, M. *Glycobiology* **2009**, *19*, 375-381.
(135) Valliere-Douglass, J. F.; Brady, L. J.; Farnsworth, C.; Pace, D.; Balland, A.; Wallace, A.; Wang, W.; Treuheit, M. J.; Yan, B. *Glycobiology* **2009**, *19*, 144-152.
(136) Wada, Y.; Tajiri, M.; Ohshima, S. *Journal of proteome research* **2010**, *9*, 1367-1373.
(137) Graham, M. E.; Thaysen-Andersen, M.; Bache, N.; Craft, G. E.; Larsen, M. R.; Packer, N. H.; Robinson, P. J. *Journal of proteome research* **2011**, *10*, 2725-2733.
(138) Halim, A.; Nilsson, J.; Ruetschi, U.; Hesse, C.; Larson, G. *Molecular & cellular proteomics : MCP* **2012**, *11*, M111.013649.
(139) Teo, C. F.; Ingale, S.; Wolfert, M. A.; Elsayed, G. A.; Not, L. G.; Chatham, J. C.; Wells, L.; Boons, G. J. *Nature chemical biology* **2010**, *6*, 338-343.
(140) Hanisch, F. G. *Analytical chemistry* **2011**, *83*, 4829-4837.
(141) Muller, S.; Goletz, S.; Packer, N.; Gooley, A.; Lawson, A. M.; Hanisch, F. G. *The Journal of biological chemistry* **1997**, *272*, 24780-24793.
(142) Seipert, R. R.; Dodds, E. D.; Lebrilla, C. B. *Journal of proteome research* **2009**, *8*, 493-501.
(143) Thakur, D.; Rejtar, T.; Karger, B. L.; Washburn, N. J.; Bosques, C. J.; Gunay, N. S.; Shriver, Z.; Venkataraman, G. *Analytical chemistry* **2009**, *81*, 8900-8907.
(144) An, H. J.; Peavy, T. R.; Hedrick, J. L.; Lebrilla, C. B. *Analytical chemistry* **2003**, *75*, 5628-5637.
(145) Clowers, B. H.; Dodds, E. D.; Seipert, R. R.; Lebrilla, C. B. *Journal of proteome research* **2007**, *6*, 4032-4040.
(146) Dodds, E. D.; Seipert, R. R.; Clowers, B. H.; German, J. B.; Lebrilla, C. B. *Journal of proteome research* **2009**, *8*, 502-512.
(147) Froehlich, J. W.; Barboza, M.; Chu, C.; Lerno, L. A., Jr.; Clowers, B. H.; Zivkovic, A. M.; German, J. B.; Lebrilla, C. B. *Analytical chemistry* **2011**, *83*, 5541-5547.

(148) Songsrirote, K.; Li, Z.; Ashford, D.; Bateman, A.; Thomas-Oates, J. *Journal of proteomics* **2010**, *73*, 1479-1490.
(149) Wuhrer, M.; Koeleman, C. A.; Hokke, C. H.; Deelder, A. M. *Analytical chemistry* **2005**, *77*, 886-894.
(150) Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom Rev* **2005**, *24*, 201-222.
(151) Zauner, G.; Deelder, A. M.; Wuhrer, M. *Electrophoresis* **2011**, *32*, 3456-3466.
(152) Hayes, C. A.; Karlsson, N. G.; Struwe, W. B.; Lisacek, F.; Rudd, P. M.; Packer, N. H.; Campbell, M. P. *Bioinformatics* **2011**, *27*, 1343-1344.
(153) Marino, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. *Nat. Chem. Biol* **2010**, *6*, 713-723.

# Chapter 3

## Suppression of peeling during the release of O-glycans by hydrazinolysis

**Radoslaw P. Kozak[1], Louise Royle[1], Richard A. Gardner[1], Daryl L. Fernandes[1] and Manfred Wuhrer [2]**

[1]Ludger Ltd., Culham Science Centre, Oxfordshire, OX14 3EB, United Kingdom

[2]Biomolecular Mass Spectrometry Unit, Department of Parasitology, Leiden University Medical Center, Postbus 9600, 2300 RC, Leiden, The Netherlands

Analytical Biochemistry (2012) 119-128.

## Abstract

The analysis of O-glycans is essential for better understanding their functions in biological processes. Although many techniques for O-glycan release have been developed, the hydrazinolysis release method is the best for producing O-glycans with free reducing termini in high yield. This release technique allows the glycans to be labelled with a fluorophore and analysed by fluorescence detection. Under the hydrazinolysis release conditions, a side reaction is observed which causes the loss of monosaccharides from the reducing terminus of the glycans (known as peeling). Using bovine fetuin (as it contains the sialylated O-glycans most commonly found on biopharmaceuticals) and bovine submaxillary glands mucin (BSM) we here demonstrate that peeling can be greatly reduced when the sample is buffer exchanged prior to hydrazinolysis with either solutions of 0.1% trifluoroacetic acid (TFA) or low molarity (100, 50, 20 and 5 mM) ethylenediaminetetraacetic acid (EDTA). Addition of calcium chloride to fetuin resulted in an increase in peeling, while subsequent washing with EDTA abolished this effect, suggesting a role of calcium and possibly other cations in causing peeling. The presented technique for sample preparation prior to hydrazinolysis greatly reduces the level of undesirable cleavage products in O-glycan analysis and increases the robustness of the method.

## Introduction

Changes in glycosylation have been associated with many states of health and diseases providing diagnostic and prognostic information.[1-3] Analysis of glycosylation is not only important in clinical and biological research but also in biopharmaceutical research where glycosylation throughout the drug life cycle must be optimised and monitored.

There are two main types of protein glycosylation found in eukaryotes, N-linked glycosylation and O-linked glycosylation. While the methods for release and analysis of N-glycans (asparagine-linked) are well established, O-glycan (covalently linked to serine or threonine) release and analysis still remains very challenging. Various enzymatic and chemical techniques for the release and recovery of O-glycans have been developed. Each of these techniques has advantages and disadvantages. An enzymatic release method would be preferred as this would avoid degradation or chemical modification. Unfortunately, no enzyme has yet been ìsolated for the universal release of O-glycans from glycoproteins. There are two O-glycanases available [4-6] but their mode of action is restricted to the liberation of the neutral core 1 disaccharide Gal-GalNAc from serine or threonine. Therefore, chemical release is currently the only effective method for obtaining the full range of O-glycans. Several techniques for chemical release of O-glycans from glycoproteins have been reported. The most common one is reductive β-elimination, a method that combines the use of sodium hydroxide and sodium borohydride. [7,8] Glycan release is accomplished by sodium hydroxide whilst sodium borohydride reduces the terminal sugar converting it to an alditol, thus preventing the oligosaccharide from degrading during and after the release. This degradation, referred to as peeling, produces the

glycan structure NeuAcα2-3Gal from glycoproteins carrying sialylated T-antigens (Neu5Acα2-3Galβ1-3GalNAc) [9] and happens when 3-O-substituents of the reducing termini of the released O-glycans 'peel' off causing loss of the terminal monosaccharide(s). Peeling is a general problem found with chemical release methods for O-glycans and results in poor repeatability and stoichiometry for analytical glycan profiles. The mechanism for formation of this degradation product is not fully understood. Due to the use of the reducing agent to prevent peeling, the reductive β-elimination method results in glycans with reduced ends. This makes it impossible to introduce a fluorescent or colorimetric label such as 2-aminobenzamide (2-AB) or 2-aminobenzoic acid (2-AA) for fluorescence detection or 1-phenyl-3methyl-5pyrazolone (PMP) for UV detection.[10] Therefore alternative detection methods such as mass spectrometry [11] or HPAEC-PAD (high pH anion exchange chromatography with pulsed amperometric detection) [12] have to be used for the analysis of alditols.

To overcome this problem with alditols non-reductive β-elimination methods have been developed. These methods use various reagents such as ammonia [13], 70% (w/v) aqueous ethylamine [14] and lithium hydroxide [15] to release O-glycans without subsequent reduction. However, these methods have not found wide acceptance in the scientific community.

A combination of enzymatic digestion with chemical release using solid-phase permethylation of O-linked oligosaccharides has been described by Goetz *et al.* [16] Firstly, glycoproteins are incubated with pronase and then the whole reaction mixture is subjected to solid phase permethylation. This technique allows the release of O-glycans with free reducing termini in high yield without peeling. However, as the

glycans are released they become permethylated including methylation of the reducing-end anomeric hydroxyl group. Therefore, they cannot be labelled with a fluorophore of chromophore and mass spectrometric analysis is required.

Recently, Zauner et al.[17], Wang et al.[18] and Furukawa et al.[19] have described a combined method for O-glycan release and labelling using a β-elimination method with dimethylamine, [17] ammonia [18] or sodium hydroxide [19] for the release and concomitant labelling with 1-phenyl-3-methyl-5-pyrazolone (PMP). The technique is quick, provides good release and labelling and is reported to have a reduced or very low degree of peeling. However, PMP labelling only allows UV detection, which is not as sensitive and specific as fluorescence detection, or mass spectrometric analysis of the released PMP-labelled glycans.

Another widely used procedure for the release of O-glycans is the hydrazinolysis method.[9,20,21] Depending on the specific protocol, N- and/or O-glycans can be released. Whilst N-glycans are released from glycoproteins using more vigorous conditions (85°C to 100°C, 5-16 hours), [21] O-glycans are released and recovered under milder conditions (60°C, 6 hours).[9] Although hydrazinolysis of O-glycans produces intact glycans with free reducing termini in high yield, a major concern is the occurrence of undesirable peeling. Merry et al.[9] reported that removal of water from the sample before hydrazinolysis is essential to minimise peeling. The presence of water may lead to the alkaline conditions that favour β-elimination and produce higher amounts of peeling. Although removal of water does reduce peeling, it still remains a significant problem that has been proved difficult to control.

Despite significant amounts of work to develop improved methods for O-glycan release and recovery, currently there are no general, fast methods available for the removal of O-glycans with free reducing termini from glycoproteins (to allow

sample analysis by fluorescent detection) in high yield without partial degradation. In the course of our studies into this peeling phenomenon, we investigated a variety of ways to reduce or completely avoid peeling. In this article we examine the effect of removing buffer salts and other low-molecular weight materials from glycoprotein samples by a range of acid and ethylenediaminetetraacetic acid (EDTA) washes prior to hydrazinolysis.

## Materials and Methods

### Materials

Anhydrous hydrazine (99.9%) and all other reagents for hydrazinolysis were from Ludger Ltd (Oxford, UK). EDTA (99.0%) was obtained from VWR (Leicestershire, UK). Acetonitrile (Romil; 190 SpS for UV/ gradient quality) was obtained from Charlton Scientific (Charlton, Oxon, UK). Centrifugal filter devices with a molecular weight cut off membrane of 10 kDa were obtained from Fisher Scientific UK (Loughborough, Leicestershire, UK). Bovine fetuin was from Ludger Ltd (Oxford, UK) and bovine submaxillary gland mucin (Type I-S) (BSM) was obtained from Sigma-Aldrich (Zwijndrecht, The Netherlands). All other reagents were obtained from Sigma-Aldrich (Dorset, UK).

### Sample preparation for hydrazinolysis

Fetuin samples (375 µg) were dissolved in 150 µL of water or a range of solvents: 0.1 M PBS; 0.1% TFA; 0.1% HCl; 0.1% $H_2SO_4$; 0.1% HCOOH; 0.1% $CH_3COOH$; 100 mM, 50 mM, 20 mM, 5 mM EDTA. Each 150 µL solution of fetuin glycoprotein was then transferred to a separate centrifugal filter device (10 kDa,

molecular weight cut off membrane), which had been pre-washed with water (5 mL). Each centrifugal filter device was centrifuged at 4000 rpm for 10-12 min (until between 0.3 and 0.5 mL of solution remained).

The washing was then repeated a further five times with 5 mL of the appropriate washing solution for each sample. The remaining 0.3 to 0.5 mL of solution was divided and transferred into three pyrolysed 5 mL glass vials (for triplicate release) and dried down for 16 hours by vacuum centrifugation (without heating) prior to addition of hydrazine.

BSM samples (375 µg) were prepared as for fetuin with solvents: water; 0.1% TFA and 100 mM EDTA.

## Release of O-glycans by hydrazinolysis

The O-glycans were released from fetuin or BSM glycoprotein by the addition of hydrazine and incubation at 60°C for 6 hours as previously described by Merry et al.,[9] Patel et al.[21] and Ashford et al.[20] Hydrazine was removed by centrifugal evaporation. The samples were placed on ice for 20 min (0°C) and were re-*N*-acetylated by the addition of a 0.1 M sodium bicarbonate solution (200 µL) and acetic anhydride (21 µL). Samples were mixed and incubated at 0°C for 10 min. A further aliquot of acetic anhydride (21 µL) was added to each sample followed by vortexing and incubation at room temperature for 60 min. Samples were cleaned up by passing them through LudgerClean CEX cartridges (Ludger Ltd). The glycans were eluted off the cartridges using water (3 x 0.5 mL). Eluates were dried by vacuum centrifugation prior to fluorescent labelling.

## Fluorescent labelling

Released O-glycans were labelled with 2-aminobenzamide (2-AB) according to Bigge et al.[22] using a Ludger 2-AB Glycan Labelling Kit. The released glycans were incubated with labelling reagents for 3 hours at 65°C. The 2-AB labelled glycans were cleaned up using LudgerClean S Cartridges (Ludger Ltd). 2-AB labelled glycans were eluted from the LudgerClean S Cartridges with water (2 x 0.5 mL). The samples were evaporated to dryness under high vacuum using centrifugal evaporation and re-suspended in water (100 µL) for further analysis.

## Exoglycosidase digestions

Exoglycosidase sequencing was performed according to Royle et al..[23] The released, 2-AB labelled glycans were incubated with exoglycosidases at standard concentrations in a final volume 10 µL in 50 mM sodium acetate for 16 hours at 37°C. The enzymes used were: Sialidase from *Arthobacter ureafaciens* (a368S): specific for α2-3, 6, 8, 9 sialic acids (E-S001; QABio, Palm Desert, CA); Beta galactosidase from *Streptococcus pneumoniae* (b4G): specific for β1-4 galactose (E-BG07; QABio, Palm Desert, CA); Bovine kidney α-fucosidase (a6F): specific for α1-6>2 fucose (EC 3.2.1.51; Sigma-Aldrich, UK). After digestion, samples were separated from the exoglycosidases by binding onto a protein-binding plate for 60 min followed by elution of the glycans from the plate with water. The samples were analysed by HILIC-HPLC.

## HPLC analysis

2-AB labelled samples were analysed by HILIC-HLPC using 4.6 x 150 mm LudgerSep-N2 column (Ludger Ltd) on a 2795 HPLC with a 2475 fluorescence detector, controlled by Empower software (Waters, Manchester, UK). Solvent A was 50 mM ammonium formate pH 4.4, Solvent B was acetonitrile. Samples were injected in 20% aqueous/80% acetonitrile; injection volume 25 µL. Gradient conditions were 20% - 38% A at flow rate of 1.0 ml/min over 36 min, followed by 2 min at 100% solvent A, then finishing with 20% solvent A, giving a total run time of 45 min. Waters GPC software with a cubic spline fit was used to allocate GU values to peaks. 2-AB labelled glucose homopolymer (CAB-GHP-30, Ludger Ltd) was used as a system suitability standard as well as an external calibration standard for GU allocation. [23]

| Acis used for washing fetuin sample prior to hydrazinolysis | Molarity (mM) | pH |
|---|---|---|
| 0.1% HCl | 3.6 | 1.4 |
| 0.1% H2SO4 | 10.2 | 1.7 |
| 0.1% HCOOH | 21.7 | 2.7 |
| 0.1% CH3COOH | 16.7 | 3.3 |
| 0.1% TFA | 8.8 | 2.0 |

**Table 1.** Molarity and pH of acid solutions used for sample cleanup prior hydrazinolysis.

## Results and Discussion

Release and recovery of O-glycans with free reducing termini is important for structural and functional analysis. Using fetuin as a model substance, we evaluated the effect of removing buffer salts and other low-molecular weight materials from glycoprotein samples before O-mode hydrazinolysis, to see if these measures could reduce peeling.

Bovine fetuin was buffer exchanged by washing with water and a range of inorganic and organic acid solutions: 0.1% TFA; 0.1% HCl; 0.1% $H_2SO_4$; 0.1% HCOOH; 0.1% $CH_3COOH$. The O-glycans were released using anhydrous hydrazine, [9] labelled with 2-aminobenzamide (2-AB) and analysed by HILIC-HPLC. To find out if the amount of degradation product was reduced when performing a solvent exchange prior to hydrazinolysis, three control experiments were performed which included using the dried fetuin glycoprotein sample without a wash step and washing a fetuin glycoprotein sample with only water, or with 0.1 M PBS.

The HPLC profiles obtained from the fetuin sample dried without cleanup prior to hydrazinolysis and from the fetuin samples cleaned up with water, 0.1 M PBS and 0.1% TFA using a centrifugal filtration device (MWCO membrane,10 kDa) were compared (Figure 1). Structures of intact O-glycans as well peeling products were assigned on the basis of previously reported GU values [9,23] and were consistent with published results.[24,25] The fetuin profile contained core 1: Galβ1-3GalNAc (peak 1), mono-sialylated core 1 O-glycans: Neu5Acα2-3Galβ1-3GalNAc (peak 3) and Neu5Acα2-6(Galβ1-3)GalNAc (peak 4); di-sialylated core 1 O-glycan: Neu5Acα2-3Galβ1-3(Neu5Acα2-6)GalNAc (peak 5); and di-sialylated core 2 O-glycan:

Neu5Acα2-3Galβ1-3(Neu5Acα2-3Galβ1-4GlcNAcβ1-6)GalNAc (peak 6). The peeled product Neu5Acα2-3Gal (peak 2, Figure 1) was also detected.

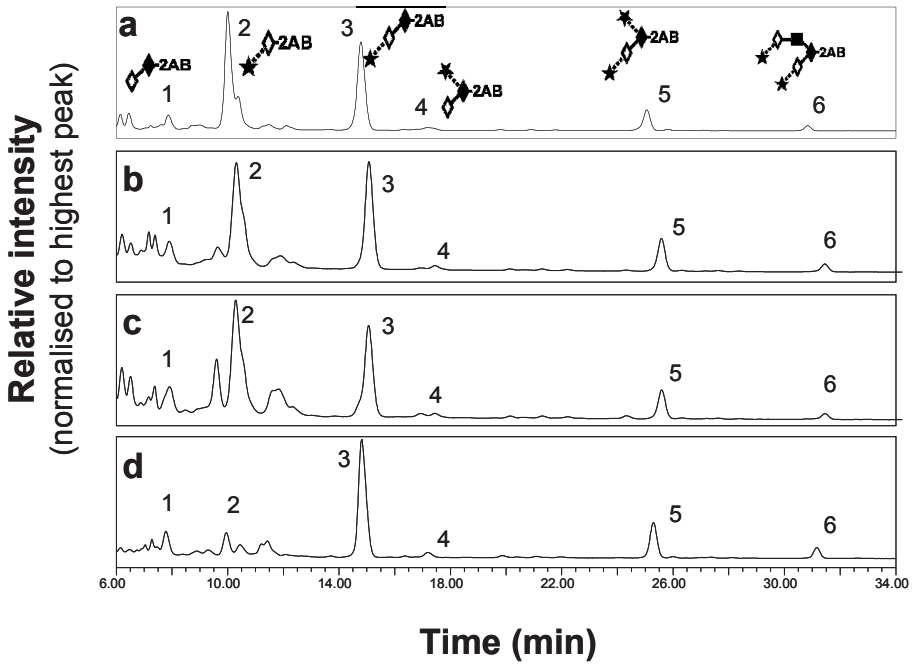



**Figure 1.** Comparison of HPLC O-glycan profiles of bovine fetuin following buffer exchange with a range of solutions, hydrazinolysis, and 2-AB labelling. The following buffer exchange procedures were applied: (a) no washing, (b), water wash, (c) 0.1 M PBS wash, (d) 100 mM EDTA wash. The O-glycans released by hydrazinolysis were 2-AB labelled and compared by HILIC-HPLC with fluorescence detection. Peak 2 is the peeled product. The following symbols are used to depict glycan structures [26]: ◊, galactose; ♦, *N*-acetylgalactosamine; ◈, fucose; ■, *N*-acetylglucosamine; ★, *N*-acetylneuraminic acid; ☆, *N*-glycolylneuraminic acid; dashed line, α-linkage; solid line, β-linkage.

The highest degree of peeling was observed for samples which were not cleaned up. Under these conditions, the peeled product Neu5Acα2-3Gal (peak 2, Figure 1a) had an average relative abundance of 58% (Table 2). The samples that were cleaned up in water or 0.1 M PBS showed a lower degree of peeling (36% or

45% relative intensity for peak 2, respectively; Figure 1b-c, Table 3). The most pronounced reduction of peeling products was apparent in the samples that were washed with 0.1% TFA. These samples showed a significantly reduced amount of peeled product (19% peak 2, Figure 1d, Table 3) as compared to the samples that were not cleaned up or were washed with water.

| Fetuin samples not cleaned up prior to hydrazinolysis | | Glycan structure | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | |
| | | Peak area [%] | | | | | |
| Experiment 1 | A | 9.7 | 68.3 | 15.0 | 0.7 | 5.4 | 1.0 |
| | B | 10.3 | 71.7 | 12.1 | 0.7 | 4.3 | 0.9 |
| | C | 9.7 | 72.0 | 12.3 | 0.5 | 4.5 | 0.9 |
| Experiment 2 | A | 4.2 | 49.9 | 34.2 | 2.2 | 7.7 | 1.8 |
| | B | 4.0 | 57.0 | 29.6 | 1.3 | 6.7 | 1.4 |
| | C | 7.7 | 30.1 | 46.8 | 2.8 | 10.1 | 2.5 |
| Average % Area | | 7.6 | **58.2** | 25.0 | 1.4 | 6.5 | 1.4 |
| STDev | | 2.6 | 14.9 | 13.0 | 0.9 | 2.0 | 0.6 |

**Table 2.** Average relative abundance of O-glycans from fetuin samples not cleaned up prior to hydrazinolysis.

Following the findings that TFA significantly reduced the occurrence of peeling we decided to investigate other inorganic and organic acids to see their effect, particularly with regard to the degree of peeling. A number of different acids were investigated. The samples were buffer exchanged by washing with: 0.1% HCl, 0.1% $H_2SO_4$, 0.1% HCOOH and 0.1% $CH_3COOH$ (Table 1). The HPLC profiles obtained from the fetuin samples cleaned up by using a centrifugal filtration device (10 kDa MWCO membrane) with the different acids were compared (Figure 2). The samples that were prepared using 0.1% HCl, 0.1% $H_2SO_4$ and 0.1% HCOOH, showed degrees of peeling which were very similar to those of the 0.1% TFA washes (Table 3). However, three of the acid-treated samples (0.1% HCl, 0.1% $H_2SO_4$ and 0.1% $CH_3COOH$) also showed an increase in de-sialylation (peak 1, Figure 2a, b and d,

Table 3). The incidence of the non-sialylated core 1 structure (peak 1, Figure 2a-d) relative to the sialylated structures (peaks 3-6, Figure 2a-d) was much higher for these acid washed samples than for samples cleaned up with water or 0.1% TFA. From these results, it is clear that although organic and inorganic acid washes, with the exception of 0.1 % $H_2SO_4$, decrease the degree of peeling, they also increase the degree of de-sialylation. De-sialylation is assumed to occur during the vacuum centrifugation step which would explain why only the highly volatile acids (HCOOH and TFA) do not cause desialylation.

| Structure | | Fetuin samples buffer exchanged by washing with: | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | no cleanup | water | 100mM PBS | 0.1% TFA | 0.1% HCl | 0.1% $H_2SO_4$ | 0.1% HCOOH | 0.1% $CH_3COOH$ | 100mM EDTA | 50mM EDTA | 20mM EDTA | 5mM EDTA |
| | **Average % area** | **7.6** | **8.2** | **8.2** | **9.2** | **13.0** | **35.4** | **8.6** | **7.5** | **7.1** | **8.0** | **6.9** | **6.5** |
| | *STDEV* | *2.6* | *1.5* | *0.1* | *2.2* | *4.8* | *7.8* | *4.5* | *1.7* | *1.2* | *0.1* | *0.4* | *0.2* |
| | p-value | - | 0.9 | 0.8 | 0.4 | 0.2 | **0.0** | 0.9 | 0.4 | 0.4 | 0.9 | 0.8 | 0.7 |
| | **Average % area** | **58.2** | **36.0** | **45.2** | **19.1** | **20.4** | **17.6** | **24.6** | **27.4** | **17.4** | **16.4** | **23.4** | **22.1** |
| | *STDEV* | *14.9* | *6.5* | *3.1* | *7.3* | *6.0* | *6.1* | *14.7* | *10.9* | *3.0* | *1.8* | *2.5* | *3.9* |
| | p-value | - | **0.0** | 0.4 | **0.0** | **0.0** | **0.0** | **0.0** | **0.0** | **0.0** | 0.2 | 0.2 | 0.3 |
| | **Average % area** | **25.0** | **40.0** | **33.8** | **49.8** | **46.9** | **31.9** | **46.5** | **42.5** | **51.8** | **54.6** | **48.7** | **48.3** |
| | *STDEV* | *13.0* | *3.6* | *2.1* | *4.3* | *4.0* | *9.0* | *10.5* | *13.1* | *2.9* | *0.5* | *1.4* | *4.8* |
| | p-value | - | **0.0** | 0.6 | **0.0** | **0.0** | 0.2 | **0.0** | **0.0** | **0.0** | 0.2 | 0.3 | 0.4 |
| | **Average % area** | **1.4** | **1.7** | **1.4** | **2.6** | **3.9** | **9.8** | **2.3** | **2.1** | **2.7** | **2.6** | **2.4** | **2.4** |
| | *STDEV* | *0.9* | *0.7* | *0.0* | *1.0* | *1.7* | *12.1* | *0.6* | *0.5* | *0.4* | *0.3* | *0.4* | *0.1* |
| | p-value | - | **0.0** | 1.0 | 0.1 | **0.0** | **0.0** | **0.0** | **0.0** | **0.0** | 0.4 | 0.5 | 0.4 |
| | **Average % area** | **6.5** | **11.3** | **9.4** | **14.1** | **12.7** | **5.9** | **14.2** | **3.0** | **16.8** | **14.9** | **14.9** | **16.4** |
| | *STDEV* | *2.0* | *1.8* | *0.8* | *2.6* | *1.3* | *1.8* | *14.2* | *3.0* | *1.7* | *2.1* | *1.2* | *0.5* |
| | p-value | - | **0.0** | 0.2 | **0.0** | **0.0** | **0.0** | **0.0** | **0.0** | **0.0** | **0.0** | 0.2 | 0.1 |
| | **Average % area** | **1.4** | **2.7** | **2.1** | **3.6** | **3.1** | **3.5** | **3.7** | **3.6** | **4.2** | **3.6** | **3.8** | **4.4** |
| | *STDEV* | *0.6* | *0.5* | *0.3* | *1.0* | *0.6* | *3.6* | *0.9* | *3.6* | *0.6* | *0.5* | *0.1* | *0.6* |
| | p-value | - | **0.0** | 0.2 | **0.0** | **0.0** | 0.3 | **0.0** | **0.0** | **0.0** | **0.0** | 0.1 | **0.0** |

**Table 3.** Comparison of the average relative abundance, standard deviation and significance level (p-value) of O-glycans from fetuin samples that had been buffer exchanged prior to hydrazinolysis. The significance level was calculated comparing the control condition (Table 2) with various treatments. P-values are given in bold for samples where changes were significant (p-value ≤ 0.05).

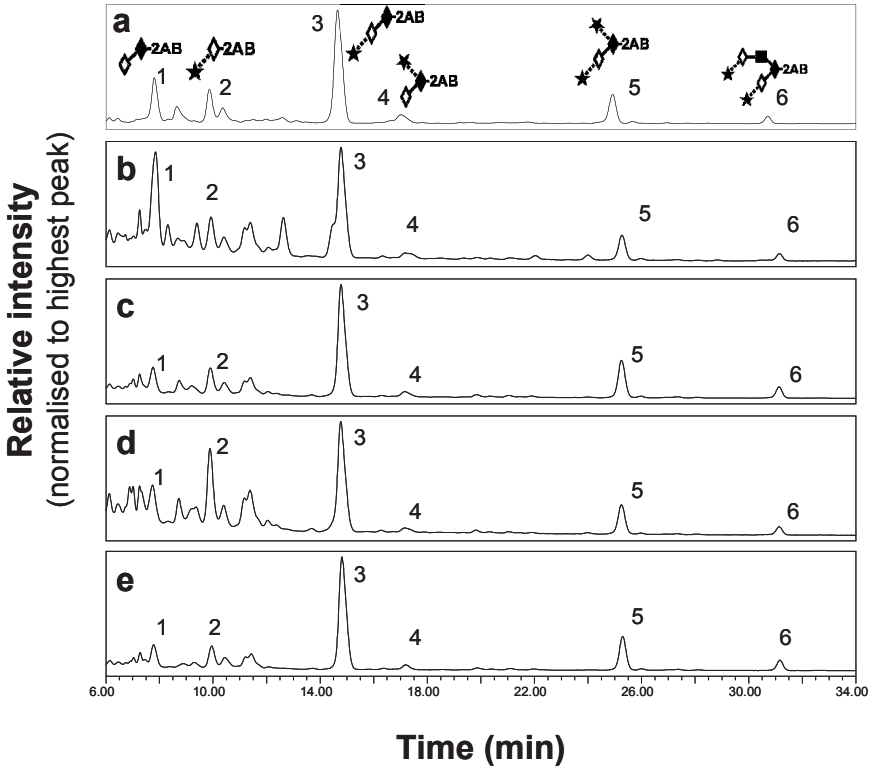



**Figure 2.** Fetuin O-glycan profiling following buffer exchange with a range of acids. The following buffer exchange procedures were applied: (a) 0.1% HCl, (b) 0.1% $H_2SO_4$, (c) 0.1% HCOOH, (d) 0.1% $CH_3COOH$, (e) 0.1% TFA. Subsequently, glycans were released by hydrazinolysis, 2-AB labelled and compared by HILIC-HPLC with fluorescence detection. Peak 2 is the peeled product.

We hypothesize that the suppression of peeling achieved by the acid washes could be due to the removal of salts that could possibly interfere with the release of glycans and stimulate the peeling reaction. We assume that the chelation of cations may promote salt removal. Therefore, we decided to study the effect of washes with solutions of the chelating agent ethylenediaminetetraacetic acid (EDTA) on the degree of peeling.

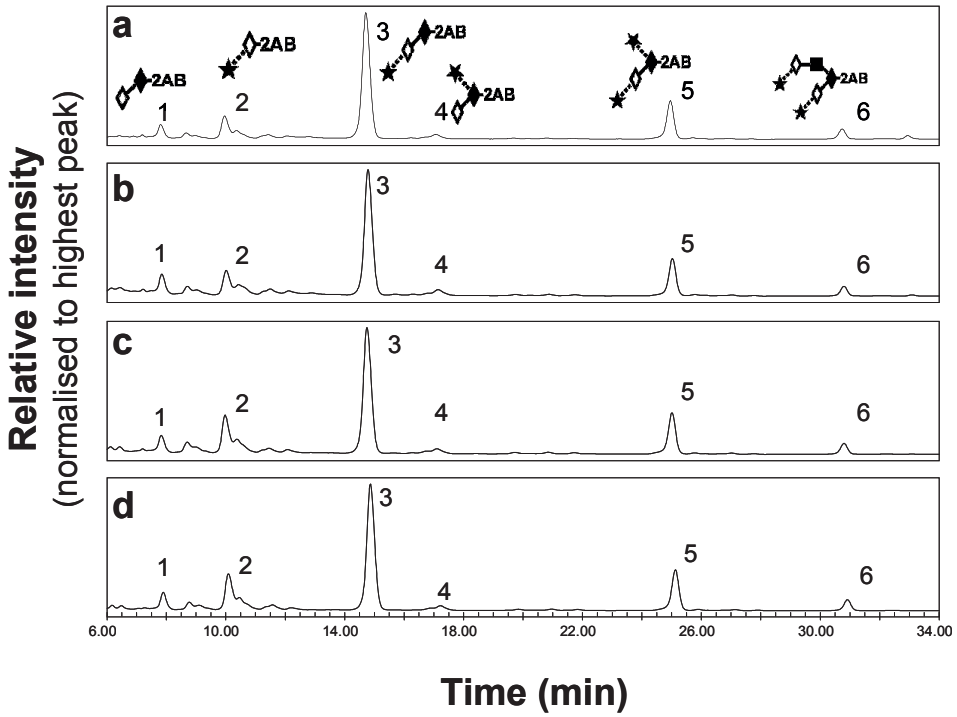



**Figure 3.** Fetuin O-glycan profiling after hydrazinolysis and buffer exchange with EDTA solutions. Sample washing was performed using (a) 100 mM EDTA, (b) 50 mM EDTA, (c) 20 mM EDTA, (d) 5 mM EDTA. Subsequently, glycans were released by hydrazinolysis, 2-AB labelled and compared by HILIC-HPLC with fluorescence detection. Peak 2 is the peeled product

Fetuin glycoprotein was cleaned up by centrifugal filtration with a range of solutions of EDTA with different concentrations (100, 50, 20 and 5 mM). The HPLC profiles (Figure 3) show that there were no significant differences in the relative abundance of peeling products between the samples washed with these different concentrations of EDTA solutions and that the relative intensity of the degradation product is similar to the 0.1% TFA cleanup (peak 2, Table 3). We therefore postulate

that the EDTA (and possibly the TFA) could be reducing peeling by removing cations, thereby preventing their interference during hydrazinolysis.

The repeatability of these cleanup procedures was tested over a 12 month period for 0.1% TFA and over a 3 month period for 100 mM EDTA. The O-glycans were released in triplicate from fetuin and labelled with 2-AB prior to HILIC-HPLC analysis. We used different batches of the fetuin glycoprotein, TFA and hydrazine to check the repeatability of the cleanup method (Figure 4). The relative amounts of released O-glycans are consistent across all the experiments and the ratio of the sialylated core 1 (peak 3, Figure 4) to its peeled product (peak 2, Figure 4) is much higher when 0.1% TFA or 100 mM EDTA were used instead of water to prepare the samples for hydrazinolysis.

In order to evaluate the effect of cations on the release of O-glycans and to demonstrate the role of EDTA washes in cation removal and suppression of peeling, a further experiment was performed. Fetuin was dissolved in 100 µL of a 100 mM $CaCl_2$ solution. Half of this fetuin solution was buffer exchanged by washing with 100 mM EDTA prior to hydrazinolysis and the other half was dried down without further manipulation. The two samples were then subjected to hydrazinolysis and O-glycan analysis. Experiments were performed in triplicate providing consistent results. The HPLC profiles showed that peeling is higher (peak 2, Figure 5a) and the yield of O-glycans lower (Figure 5a) for the sample that had not been buffer exchanged after addition of $CaCl_2$ compared to the sample that had been buffer exchanged with 100 mM EDTA (Figure 5b). This experiment showed that the presence of calcium chloride interferes with the O-glycan release and that the removal of calcium prior to hydrazinolysis reduces peeling.

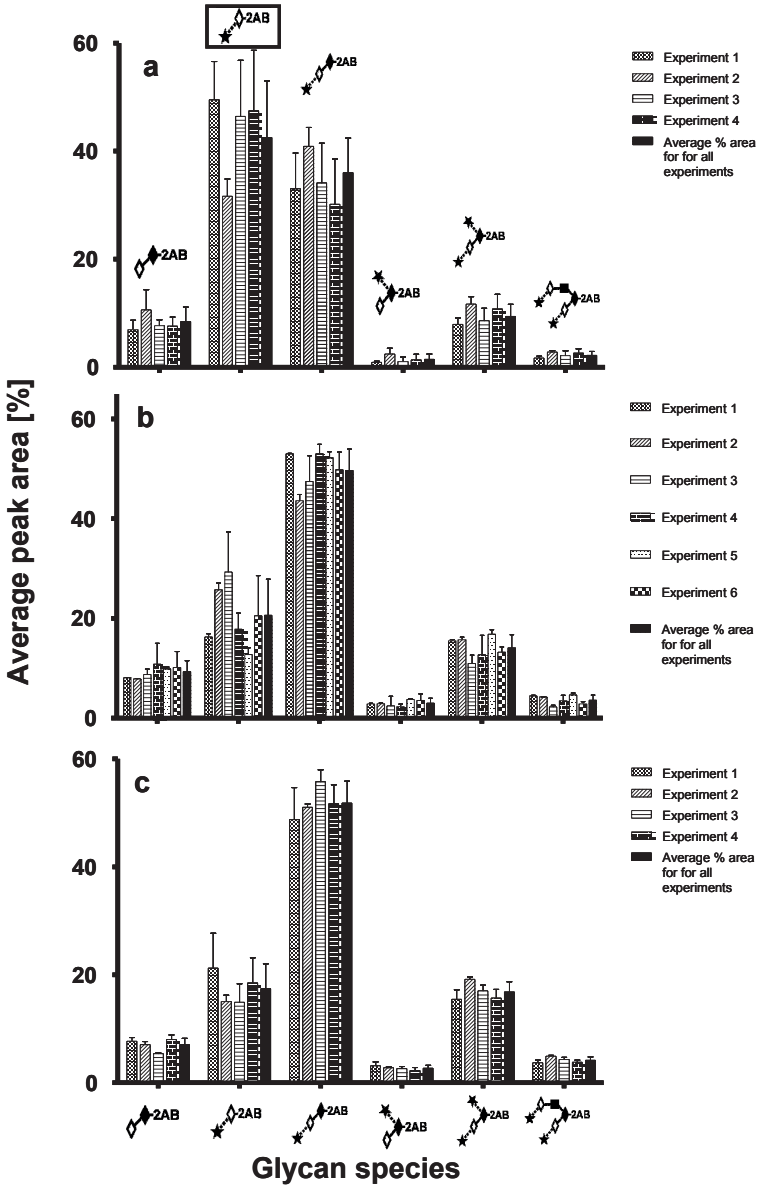



**Figure 4.** Comparison of fetuin O-glycan patterns obtained after applying various washing conditions. Samples were washed using (a) water, (b) 0.1% TFA, (c) 100 mM EDTA, followed by hydrazinolysis, 2-AB-labelling and HILIC-HPLC analysis with fluorescence detection. The average abundance of O-glycans was determined from experiments performed over a 3 month period for water, a 12 month period for 0.1% TFA and 3 month period for 100 mM EDTA.

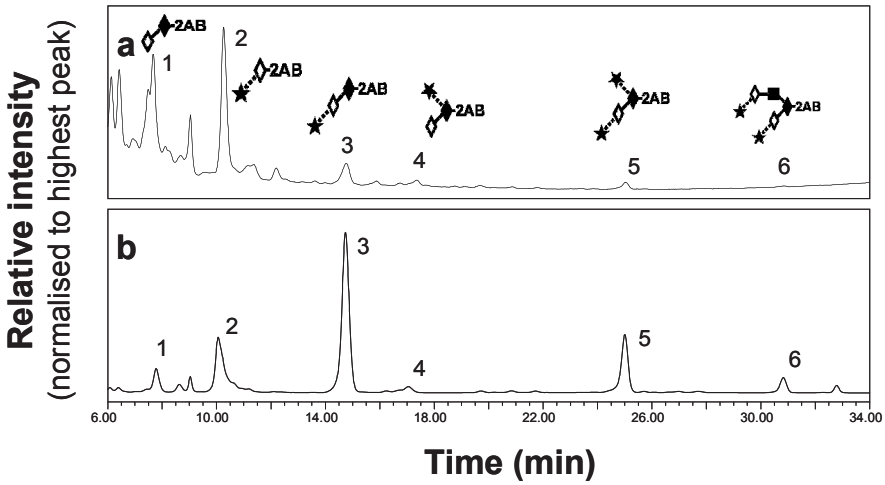



**Figure 5.** O-glycan peeling caused by the addition of calcium chloride may be prohibited by EDTA washes. Fetuin samples were dissolved in 100 mM CaCl2 and subjected to hydrazinolysis either directly (a) or after buffer exchange with 100 mM EDTA (b) followed by 2-AB-labelling and HILIC-HPLC analysis with fluorescence detection. Peak 2 is the peeled product.

Both the 0.1% TFA and 100 mM EDTA methods were also tested on bovine submaxillary gland mucin (BSM) as the O-glycans from mucins are also prone to peeling. Mucin samples were cleaned by centrifugal filtration with 0.1% TFA or 100 mM EDTA. The BSM O-glycans were released using anhydrous hydrazine,[9] labelled with 2-aminobenzamide (2-AB) and analysed by HILIC-HPLC.

The HPLC profiles obtained from the BSM sample without cleanup prior to hydrazinolysis and from the BSM samples cleaned up with water, 0.1% TFA or 100 mM EDTA using a centrifugal filtration device (MWCO membrane,10 kDa) were compared (Fig. 6). The profiles of the O-glycan pools from BSM show 11 peaks (Figure 6).

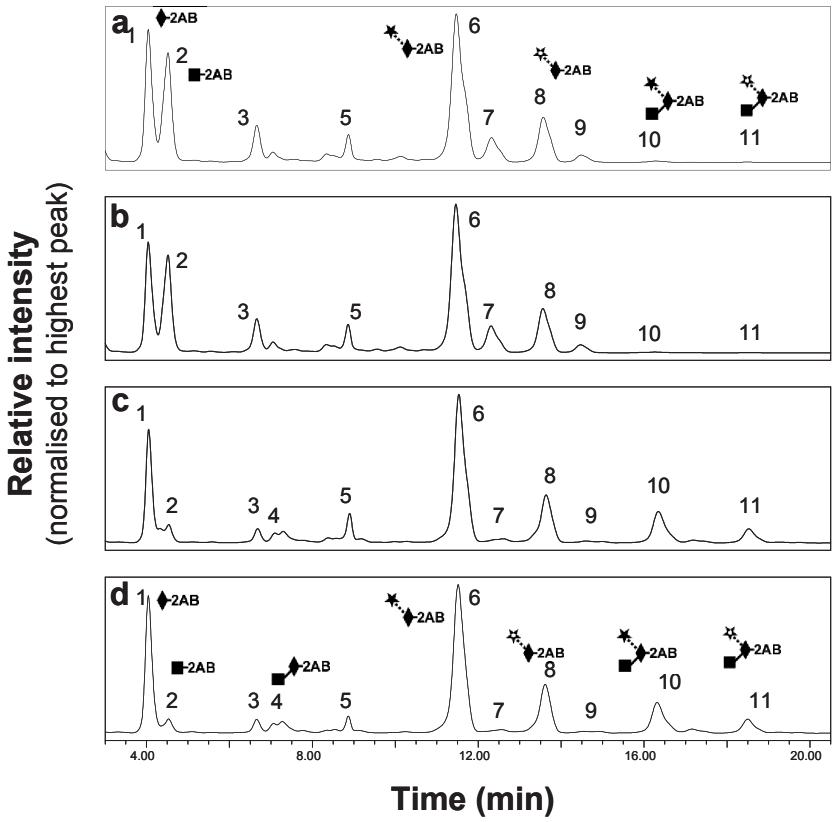



**Figure 6.** Comparison of HPLC O-glycan profiles of bovine submaxillary mucin following buffer exchange with a range of solutions, hydrazinolysis, and 2-AB labelling. The following buffer exchange procedures were applied: (a) no washing, (b), water wash, (c) 0.1% TFA wash, (d) 100 mM EDTA wash. The O-glycans released by hydrazinolysis from BSM were 2-AB labelled and compared by HILIC-HPLC with fluorescence detection. Peak 2 is the peeled product.

In order to characterise unknown mucin O-glycan structures, information from published LC/MS/MS analyses [17,27] was combined with the results of exoglycosidase treatments as determined by HILIC-HPLC (Figure 7). Initial digestion with sialidase showed that peaks 5 – 11 contained sialic acids (Neu5Ac or Neu5Gc) by the change in their elution position (Figures 7a-b). Digestion with bovine kidney α-fucosidase

which preferably cleaves α1-6 linked fucose showed that peak 3 contained fucose (Figures 7a and 7d). Digestion with a beta galactosidase specific for β1-4 galactose did not show any changes (Figures 7a and 7c). Six different glycan structures were identified: GalNAc (peak 1), GlcNAcβ1-3GalNAc (peak 3), Neu5Acα2-6GalNAc (peak 6); Neu5Gcα2-6GalNAc (peak 8); GlcNAcβ1-3(Neu5Acα2-6)GalNAc (peak 10) and GlcNAcβ1-3(Neu5Gcα2-6)GalNAc (peak 11) along with some peeled product GlcNAc (peak 2). BSM was found to contain also other sialylated structures (peaks 5, 7 and 9) and a fucosylated structure (peak 3) which were not fully structurally elucidated.

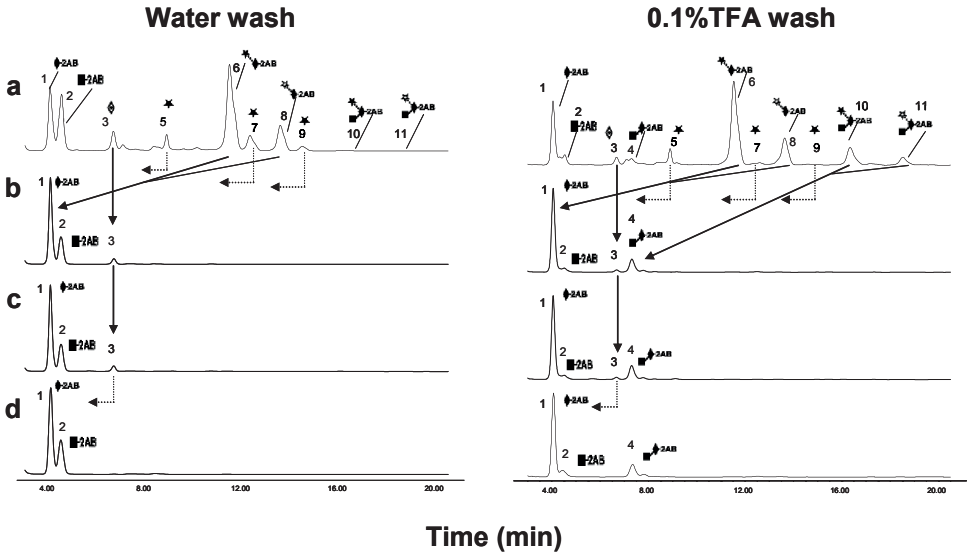



**Figure 7.** Glycan sequencing of BSM O-glycans before and after digestion with exoglycosidases. Samples were subjected to water wash (left) or 0.1% TFA wash (right) followed by hydrazinolysis and 2-AB labelling. Aliquots of the total 2AB-labelled O-glycan pool were incubated with different exoglycosidases, as shown in each panel, (a) before digestion, (b) sialidase, (c) sialidase + β-galactosidase, (d) sialidase + α-fucosidase. Following digestion, the products were analysed by HILIC-HPLC. Arrows indicate digestion pathways.

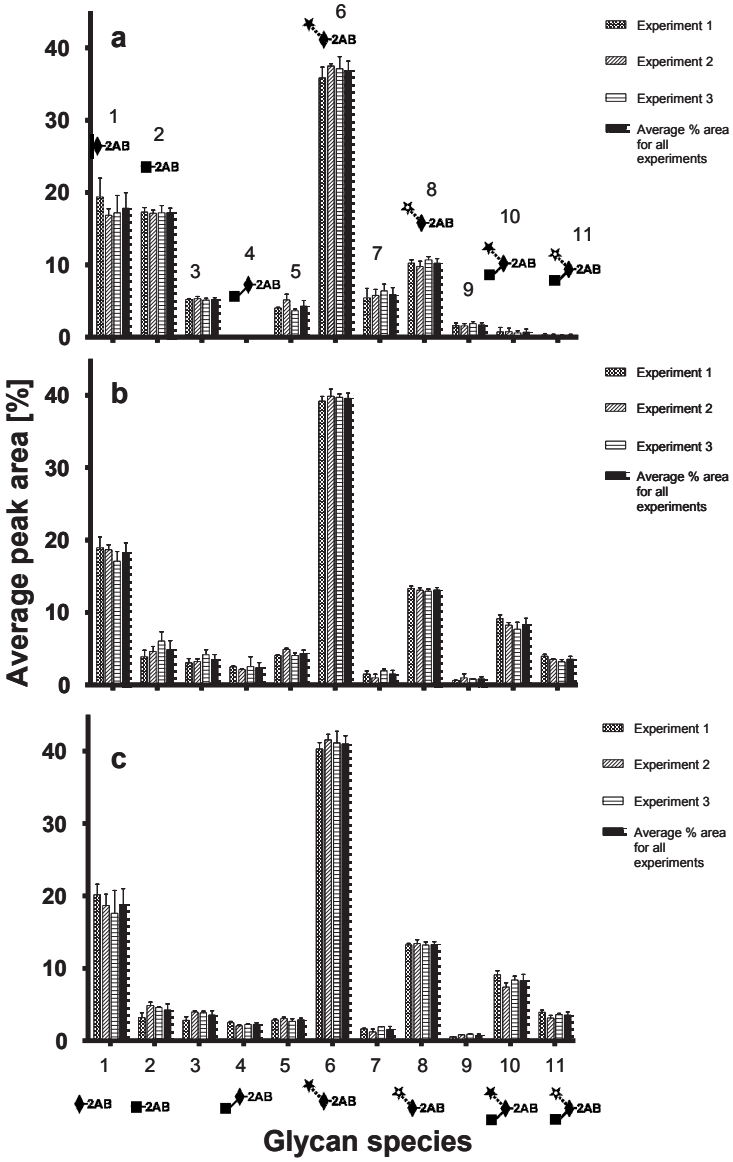



**Figure 8.** Comparison of BSM O-glycan patterns obtained after applying various washing conditions. Samples were washed using (a) water, (b) 0.1% TFA, (c) 100 mM EDTA, followed by hydrazinolysis, 2-AB-labelling and HILIC-HPLC analysis with fluorescence detection. The average abundance of O-glycans was determined from experiments performed over a 2 month period.

| BSM samples not cleaned up prior to hydrazinolysis | | Glycan structure | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ♦-2AB | ■-2AB | fucosylated structure | ■♦-2AB | sialylated structure | ✶♦-2AB | sialylated structure | ✶✶♦-2AB | sialylated structure | ✶♦■-2AB | ✶✶♦■-2AB |
| | | Peak area [%] | | | | | | | | | | |
| Experiment 1 | A | 18.2 | 17.9 | 5.5 | - | 3.8 | 35.6 | 5.9 | 10.4 | 1.9 | 0.5 | 0.2 |
| | B | 20.0 | 17.6 | 5.3 | - | 4.0 | 35.3 | 5.7 | 9.9 | 1.6 | 0.4 | 0.1 |
| | C | 17.7 | 16.5 | 5.2 | - | 4.1 | 37.2 | 6.6 | 10.5 | 1.9 | 0.2 | 0.1 |
| Experiment 2 | A | 18.7 | 17.8 | 5.1 | - | 3.7 | 36.1 | 5.8 | 10.5 | 1.7 | 0.4 | 0.2 |
| | B | 18.7 | 17.3 | 4.9 | - | 4.3 | 36.6 | 5.9 | 10.2 | 1.7 | 0.3 | 0.1 |
| | C | 17.1 | 17.4 | 5.2 | - | 4.1 | 36.9 | 6.6 | 10.5 | 1.9 | 0.3 | 0.1 |
| Average % Area | | 18.4 | 17.4 | 5.2 | - | 4.0 | 36.3 | 6.1 | 10.3 | 1.8 | 0.4 | 0.1 |
| STDev | | 0.9 | 0.5 | 0.2 | - | 0.2 | 0.7 | 0.4 | 0.2 | 0.1 | 0.1 | 0.0 |

**Table 4.** Average relative abundance of O-glycans from BSM samples not cleaned up prior to hydrazinolysis.

A large difference in the relative abundance of peeling products was apparent between the different sample clean ups. The highest occurrence of peeling was observed for samples which were not cleaned up (17% peak 2, Figure 6a, Table 5) and for samples washed with water (17%, peak 2, Figure 6b, Table 5). For the samples washed with 0.1% TFA (5%, peak 2, Figure 6c, Table 5) or 100 mM EDTA (4% peak 2, Figure 6d, Table 5) there was much less peeled product. The samples which were not cleaned up or cleaned with water also showed a significantly reduced amount of GlcNAcβ1-3(Neu5Acα2-6)GalNAc (peak 10, Figure 6a) and GlcNAcβ1-3(Neu5Gcα2-6)GalNAc (peak 11, Figure 5a) when compared to the samples that were cleaned up with 0.1% TFA or 100 mM EDTA.(Figures 6c-d).

Alongside the comparisons of the relative abundance of O-glycans present in both fetuin glycoprotein and BSM glycoprotein (Tables 3 and 5), the absolute abundance of O-glycans was also compared by making use of the quantitative response of the HPLC fluorescence detector (Supplementary Tables 1 and 2). These results showed that there was no major difference in the overall yield of O-glycans from fetuin or BSM from either the acid-washed or EDTA-washed samples when compared to the control sample (no cleanup). This indicated that there were no

major glycoprotein losses occurring during the sample washing step using the centrifugal filtration device (10 kDa MWCO membrane).

| Structure | | BSM samples buffer exchanged by washing with: | | | |
|---|---|---|---|---|---|
| | | no cleanup | water | 0.1% TFA | 100mM EDTA |
| | **Average % area** | **18.4** | **17.8** | **18.2** | **18.8** |
| | *STDEV* | *0.9* | *2.2* | *1.4* | *2.2* |
| | p-value | - | 0.5 | 0.8 | 0.5 |
|  | **Average % area** | **17.4** | **17.2** | **4.8** | **4.2** |
| | *STDEV* | *0.5* | *0.6* | *1.3* | *0.9* |
| | p-value | - | 0.7 | **0.0** | **0.0** |
| **fucosylated structure** | **Average % area** | **5.2** | **5.2** | **0.2** | **3.5** |
| | *STDEV* | *0.2* | *5.2* | *0.2* | *0.6* |
| | p-value | - | 0.8 | **0.0** | 0.1 |
| | **Average % area** | **-** | **-** | **2.4** | **2.3** |
| | *STDEV* | - | - | *0.7* | *0.2* |
| | p-value | - | - | **0.0** | **0.0** |
| **sialylated structure** | **Average % area** | **4.0** | **4.3** | **4.3** | **2.9** |
| | *STDEV* | *0.2* | *0.8* | *0.5* | *0.3* |
| | p-value | - | 0.6 | 0.4 | **0.0** |
| | **Average % area** | **36.3** | **36.8** | **39.5** | **41.0** |
| | *STDEV* | *0.7* | *1.4* | *0.7* | *1.2* |
| | p-value | - | 0.4 | **0.0** | **0.0** |
| **sialylated structure** | **Average % area** | **6.1** | **5.9** | **1.5** | **1.6** |
| | *STDEV* | *0.4* | *1.0* | *0.5* | *0.4* |
| | p-value | - | **0.0** | **0.0** | **0.0** |
| | **Average % area** | **10.3** | **10.2** | **13.1** | **13.3** |
| | *STDEV* | *0.2* | *0.6* | *0.3* | *0.4* |
| | p-value | - | 0.9 | **0.0** | **0.0** |
| **sialylated structure** | **Average % area** | **1.8** | **1.7** | **0.8** | **0.7** |
| | *STDEV* | *0.1* | *0.3* | *0.3* | *0.2* |
| | p-value | - | **0.0** | **0.0** | **0.0** |
| | **Average % area** | **0.4** | **0.7** | **8.3** | **8.3** |
| | *STDEV* | *0.1* | *0.4* | *0.9* | *0.9* |
| | p-value | - | **0.0** | **0.0** | **0.0** |
| | **Average % area** | **0.1** | **0.2** | **3.5** | **3.5** |
| | *STDEV* | *0.0* | *0.2* | *0.4* | *0.4* |
| | p-value | - | **0.0** | **0.0** | **0.0** |

**Table 5.** Comparison of the average relative abundance, standard deviation and significance level (p-value) of O-glycans from BSM samples that had been buffer exchanged prior to hydrazinolysis. The significance level was calculated comparing the control condition (Table 4) with various treatments. P-values are given in bold for samples where changes were significant (p-value ≤ 0.05).

These studies show that buffer exchange into 0.1% TFA or 100 mM EDTA prior to hydrazinolysis significantly reduces the amount of undesirable peeling. Peeled product went down from 58% for the fetuin sample without cleanup to <20% for the fetuin samples washed with 0.1% TFA or EDTA and from 17% for the BSM sample without cleanup to <5% for the BSM samples washed with 0.1% TFA or EDTA. As TFA is volatile it will be removed during the vacuum centrifugation step that is applied prior to hydrazinolysis, while EDTA will stay within the sample. The two sample cleanup methods used, 0.1% TFA and 100 mM EDTA washes, appear to work equally well. Both of them result in less peeling, more robust O-glycan profiles and, therefore, greatly improve O-glycan analysis employing hydrazinolysis.

## Acknowledgement

The authors thank Gerhild Zauner for critically reading the manuscript.

## Reference List:


(1) Ohtsubo, K.; Marth, J. D. *Cell* **2006**, *126*, 855-867.

(2) Varki, A.; Marth, J. *seminars in DEVELOPMENTAL BIOLOGY* **1995**, *6*, 127-138.

(3) Arnold, J. N.; Wormald, M. R.; Sim, R. B.; Rudd, P. M.; Dwek, R. A. *Annu. Rev. Immunol* **2007**, *25*, 21-50.

(4) Geyer, H.; Geyer, R. *Biochim. Biophys. Acta* **2006**, *1764*, 1853-1869.

(5) Iwase, H.; Hotta, K. *Methods Mol. Biol* **1993**, *14*, 151-159.

(6) Dwek, R. A.; Edge, C. J.; Harvey, D. J.; Wormald, M. R.; Parekh, R. B. *Annu. Rev. Biochem* **1993**, *62*, 65-100.

(7) Schulz, B. L.; Packer, N. H.; Karlsson, N. G. *Anal. Chem* **2002**, *74*, 6088-6097.

(8) Carlson, D. M. *J. Biol. Chem* **1966**, *241*, 2984-2986.

(9) Merry, A. H.; Neville, D. C.; Royle, L.; Matthews, B.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M. *Anal. Biochem* **2002**, *304*, 91-99.

(10) Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M. *Anal. Bioanal. Chem* **2010**, *397*, 3457-3481.

(11) Karlsson, N. G.; Wilson, N. L.; Wirth, H. J.; Dawes, P.; Joshi, H.; Packer, N. H. *Rapid Commun. Mass Spectrom* **2004**, *18*, 2282-2292.

(12) Stadheim, T. A.; Li, H.; Kett, W.; Burnina, I. N.; Gerngross, T. U. *Nat. Protoc* **2008**, *3*, 1026-1031.

(13) Huang, Y.; Mechref, Y.; Novotny, M. V. *Anal. Chem* **2001**, *73*, 6063-6069.

(14) Chai, W.; Feizi, T.; Yuen, C. T.; Lawson, A. M. *Glycobiology* **1997**, *7*, 861-872.

(15) Yamada, K.; Hyodo, S.; Kinoshita, M.; Hayakawa, T.; Kakehi, K. *Anal. Chem* **2010**, *82*, 7436-7443.

(16) Goetz, J. A.; Novotny, M. V.; Mechref, Y. *Anal. Chem* **2009**, *81*, 9546-9552.

(17) Zauner, G.; Koeleman, C. A.; Deelder, A. M.; Wuhrer, M. *Biochim. Biophys. Acta* **2012**, *1820*, 1420-1428.

(18) Wang, C.; Fan, W.; Zhang, P.; Wang, Z.; Huang, L. *Proteomics* **2011**, *11*, 4229-4242.

(19) Furukawa, J.; Fujitani, N.; Araki, K.; Takegawa, Y.; Kodama, K.; Shinohara, Y. *Anal. Chem* **2011**, *83*, 9060-9067.

(20) Ashford, D. A.; Alafi, C. D.; Gamble, V. M.; Mackay, D. J.; Rademacher, T. W.; Williams, P. J.; Dwek, R. A.; Barclay, A. N.; Davis, S. J.; Somoza, C.; . *J. Biol. Chem* **1993**, *268*, 3260-3267.

(21) Patel, T.; Bruce, J.; Merry, A.; Bigge, C.; Wormald, M.; Jaques, A.; Parekh, R. *Biochemistry* **1993**, *32*, 679-693.

(22) Bigge, J. C.; Patel, T. P.; Bruce, J. A.; Goulding, P. N.; Charles, S. M.; Parekh, R. B. *Anal. Biochem* **1995**, *230*, 229-238.

(23) Royle, L.; Mattu, T. S.; Hart, E.; Langridge, J. I.; Merry, A. H.; Murphy, N.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M. *Anal. Biochem* **2002**, *304*, 70-90.

(24) Spiro, R. G.; Bhoyroo, V. D. *J. Biol. Chem* **1974**, *249*, 5704-5717.

(25) Takasaki, S.; Kobata, A. *Biochemistry* **1986**, *25*, 5709-5715.

(26) Harvey, D. J.; Merry, A. H.; Royle, L.; Campbell, M. P.; Dwek, R. A.; Rudd, P. M. *Proteomics* **2009**, *9*, 3796-3801.

(27) Robbe, C.; Capon, C.; Flahaut, C.; Michalski, J. C. *Electrophoresis* **2003**, *24*, 611-621.

# Supplementary information

| Structure | | Fetuin samples buffer exchanged by washing with: | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | no cleanup | water | 100mM EDTA | 0.1% TFA | 0.1% HCl | 0.1% $H_2SO_4$ | 0.1% HCOOH | 0.1% $CH_3COOH$ |
| | Absolute average area (/1000000) | 2.4 | 1.5 | 3.4 | 2.8 | 3.5 | 3.9 | 2.5 | 2.4 |
| | Absolute average area (/1000000) | 19.5 | 9.1 | 8.8 | 7.4 | 5.1 | 2.6 | 7.4 | 7.7 |
| | Absolute average area (/1000000) | 14.8 | 7.9 | 24.5 | 15.2 | 12.5 | 3.5 | 17.1 | 16.4 |
| | Absolute average area (/1000000) | 0.9 | 0.3 | 1.3 | 0.8 | 1.2 | 0.4 | 0.8 | 0.7 |
| | Absolute average area (/1000000) | 3.4 | 2.1 | 8.2 | 4.0 | 3.3 | 0.9 | 5.0 | 4.8 |
| | Absolute average area (/1000000) | 0.8 | 0.5 | 2.1 | 0.9 | 0.8 | 0.2 | 1.3 | 1.3 |

**Supplementary Table 1.** Comparison of the average absolute area under the curve of O-glycans from fetuin samples that had been buffer exchanged prior to hydrazinolysis. The corresponding relative abundances are presented in Table 3.

| Structure | | BSM samples buffer exchanged by | | | |
|---|---|---|---|---|---|
| | | no cleanup | water | 0.1% TFA | 100mM EDTA |
| | Absolute average area (/1000000) | 402.4 | 262.2 | 233.9 | 221.8 |
| | Absolute average area (/1000000) | 375.0 | 259.4 | 60.5 | 50.5 |
| fucosylated structure | Absolute average area (/1000000) | 113.7 | 79.3 | 43.6 | 42.7 |
| | Absolute average area (/1000000) | 0.0 | 0.0 | 30.1 | 27.0 |
| sialylated structure | Absolute average area (/1000000) | 88.6 | 64.6 | 55.9 | 34.9 |
| | Absolute average area (/1000000) | 802.8 | 558.5 | 508.8 | 495.4 |
| sialylated structure | Absolute average area (/1000000) | 139.7 | 92.2 | 18.7 | 17.9 |
| | Absolute average area (/1000000) | 226.3 | 155.3 | 168.5 | 160.8 |
| sialylated structure | Absolute average area (/1000000) | 40.8 | 26.5 | 10.3 | 8.9 |
| | Absolute average area (/1000000) | 7.5 | 8.8 | 108.3 | 99.6 |
| | Absolute average area (/1000000) | 2.6 | 8.8 | 45.9 | 42.5 |

**Supplementary Table 2.** Comparison of the average area under the curve of O-glycans from BSM samples that had been buffer exchanged prior to hydrazinolysis. The corresponding relative abundances are presented in Table 5.

# Chapter 4

## Improved nonreductive O-glycan release by hydrazinolysis with ethylenediaminetetraacetic acid addition

**Radoslaw P. Kozak[1], Louise Royle[1], Richard A. Gardner[1], Albert Bondt[2,3], Daryl L. Fernandes[1] and Manfred Wuhrer [2,4]**

[1]Ludger Ltd., Culham Science Centre, Oxfordshire, OX14 3EB, United Kingdom

[2]Center for Proteomics and Metabolomics, Leiden University Medical Center, The Netherlands

[3]Department of Rheumatology, Erasmus University Medical Center, 3000 CA Rotterdam, The Netherlands

[4]Division of BioAnalytical Chemistry, Department of Chemistry and Pharmaceutical Sciences, VU University Amsterdam, The Netherlands

Analytical Biochemistry (2014) 29-37.

## Abstract

The study of protein O-glycosylation is receiving increasing attention in biological, medical and biopharmaceutical research. Improved techniques are required to allow reproducible and quantitative analysis of O-glycans. An established approach for O-glycan analysis relies on their chemical release in high yield by hydrazinolysis, followed by fluorescent labelling at the reducing terminus and HPLC profiling. However, an unwanted degradation known as “peeling” often compromises hydrazinolysis for O-glycan analysis. Here we address this problem using low-molarity solutions of ethylenediaminetetraacetic acid (EDTA) in hydrazine for O-glycan release. O-linked glycans from a range of different glycoproteins were analysed including bovine fetuin, bovine submaxillary gland mucin and serum IgA. The data for the O-glycans released by hydrazine with anhydrous EDTA or disodium salt dihydrate EDTA show high yields of the native O-glycans compared to the peeled product, resulting in a markedly increased robustness of the O-glycan profiling method. The presented method for O-glycan release demonstrates significant reduction in peeling and reduces the number of sample handling steps prior to release.

## Introduction

Protein glycosylation is one of the most common post-translational modifications (PTMs) that influences many protein functions.[1,2] Detailed knowledge of O-linked glycan (serine- or threonine-linked oligosaccharides) structures is invaluable for better understanding their biological functions. O-glycans have a number of different functions including, but not limited to, the protection of the underlying protein as well as epithelial cell surfaces, maintaining protein conformations and controlling epitopes and antigenicity.[3] They have been reported to determine the cell surface expression and function of cell surface receptors and may be involved in growth regulation.[3] They are also known to be involved in blood clotting, embryogenesis, development and cell death [3-5] and they are associated with cancer [6,7] and other human diseases and disorders.[8-10]

Monitoring and controlling O-glycosylation is important in biopharmaceutical development and quality control (QC) in the context of existing and potential drugs that are O-glycosylated (e.g. human erythropoietin (EPO)).[11,12] The monitoring of any changes in drug glycosylation during manufacturing processes is tightly regulated by the health authorities around the world (e.g. US Food and Drug Administration or European Medicines Agency). Biopharmaceutical manufacturers must demonstrate that any process changes do not alter the clinical safety or efficacy of the biopharmaceutical.[13,14]

A number of techniques for the release, recovery and analysis of O-glycans are available.[15] Almost all of these techniques use chemical release methods, as there is no enzyme available for universal O-glycan release from proteins. Two endo-α-*N*-acetylgalactosaminidases have been described that have a high specificity but are limited to the release of the neutral core 1 disaccharide, Galβ1-3GalNAcα1.[16-18]

The most common method for the chemical liberation of O-glycans is reductive β-elimination.[19-21] This method of release leads to O-glycans that are present in their reduced forms (alditols) which means they are not amenable to labelling with a fluorescent or UV tag. Therefore, the range of analytical methods available for the characterization of O-glycans released by reductive β-elimination is restricted, with high-pH anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) [22], mass spectrometry and NMR being the predominantly applied methods.[23-26]

Several chemical methods for obtaining O-glycans in their non-reduced form have been described, using ethylamine [27], ammonia [28] or lithium hydroxide in an automatic setup [29] but the yields of released O-glycans were low and variable. It has been reported that the best method for universal removal of O-glycans in their non-reduced form in high yields is hydrazinolysis.[15,30-32] Release of glycans by non-reductive methods is compatible with reducing-end labelling using reagents such as 2-aminobenzamide (2-AB) or 2-aminobenzoic acid (2-AA) allowing HPLC with fluorescent or colorimetric (UV) detection.[33]

Analysis of O-glycosylation by HPLC based methods is preferred for a number of reasons. For example, the HPLC or UPLC-profiling of 2-AB labelled O-glycans is quantitative, highly sensitive, has a reasonably high throughput and has the ability to provide detailed information about both monosaccharide sequence and the types of glycan linkage.[34,35]

Unfortunately, all of the chemical O-glycan release methods (reductive and non-reductive) show stepwise degradation of the polysaccharides. The degradation is observed as the removal of one sugar residue at a time from the reducing end. This unwanted side reaction is known as “peeling”. [15,32]

Peeling is a general problem when performing the release of O-glycans and often results in poor repeatability with variable amounts of the small peeled glycans. This is a major problem for comparability studies or quality control. Many researchers have investigated this peeling phenomenon and it has been reported that the degree of peeling can be minimized by removal of water prior to hydrazinolysis.[32] We have recently shown that the degree of peeling can be greatly reduced by the removal of calcium or other cations prior to hydrazinolysis by employing several washes of the glycoprotein samples with a solution of 0.1% trifluoroacetic acid (TFA) or low-molarity solutions of ethylenediaminetetraacetic acid (EDTA).[35]

Although these “washing” steps are effective, they are also time consuming and can lead to loss of sample. Here we present an improved hydrazinolysis protocol that allows efficient suppression of peeling, but does not require the series of washing steps prior to the release reaction.[35] The method relies on including EDTA – either as its disodium dihydrate salt or in its anhydrous form - as part of the hydrazinolysis reaction mixture. Analysis of the O-glycans by both HPLC with fluorescence detection and LC-MS, demonstrates the efficient suppression of peeling reactions, allowing the acquisition of O-glycan profiles in a reproducible manner. The method is shown to work on commercially available standards, but also proves to be successful when analysing IgA purified by affinity enrichment from human plasma.

# Materials and Methods

## Materials

Anhydrous hydrazine (99.9%) and all other reagents for hydrazinolysis were from Ludger Ltd (Oxford, UK). Ethylenediaminetetraacetic acid disodium dihydrate salt (EDTA salt) and anhydrous EDTA were obtained from Sigma-Aldrich (Dorset, UK). Acetonitrile (Romil; 190 SpS for UV/ gradient quality) was obtained from Charlton Scientific (Charlton, Oxon, UK). Bovine fetuin was from Ludger Ltd and bovine submaxillary gland mucin (Type I-S) (BSM) as well as secretory IgA were obtained from Sigma-Aldrich.

## Purification of serum IgA

The immunoglobulin A (IgA) was captured from 10 µL human plasma of a presumably healthy volunteer using CaptureSelectTM IgA Affinity Matrix (Life Technologies/Invitrogen) in a 96-well plate format. Captured IgAs were washed with PBS buffer and water (3 x 200 µL) and eluted with 100 mM formic acid (100 µL; Fluka, Steinheim, Germany). Eluates from 4 wells were pooled and dried by vacuum centrifugation. An estimated total of 20 µg IgA was captured per well.

## Release of O-glycans

The O-glycans were released from fetuin glycoprotein by the addition of a range of solutions: anhydrous hydrazine; 50 mM, 20 mM and 10 mM EDTA disodium dihydrate salt (EDTA salt) in hydrazine or 50 mM, 20 mM, 10 mM anhydrous EDTA in hydrazine. Samples were incubated at 60°C for 6 h.[32-34] Excess hydrazine was removed by centrifugal evaporation. The samples were placed on ice for 20 min (0°C) and were re-*N*-acetylated by the addition of a 0.1 M sodium bicarbonate

solution (200 µL) and acetic anhydride (21 µL). Samples were cleaned up by passing them through LudgerClean CEX cartridges (Ludger Ltd). The glycans were eluted off the cartridges using water (3 x 0.5 mL). Eluates were dried by vacuum centrifugation prior to fluorescent labelling.

The O-glycans from BSM and serum IgA samples were released as for fetuin using the following solutions: anhydrous hydrazine, 50 mM EDTA salt in hydrazine or 20 mM EDTA anhydrous in hydrazine.

## 2-AB labelling and purification

Released O-glycans were fluorescently labelled with 2-aminobenzamide (2-AB) by reductive amination according to the method of Bigge *et al.* [36] using a Ludger 2-AB Glycan Labelling Kit (Ludger Ltd). The released glycans were incubated with labelling reagents for 3 h at 65°C. The 2-AB labelled glycans were cleaned up using LudgerClean T1 Cartridges (Ludger Ltd). 2-AB labelled glycans were eluted from the LudgerClean T1 Cartridges with water (1 mL). The samples were dried by vacuum centrifugation and re-suspended in water (100 µL) for further analysis.

## Sialidase treatments

Sialidase digestion was performed on the released 2-AB labelled glycans in 10 µL solution containing enzyme at standard concentrations in 50 mM sodium acetate for 16 h at 37°C according to Royle *et al.*[37] The enzyme used was: sialidase from *Arthobacter ureafaciens* (a368S) which is specific for α2-3, α2-6, α2-8 and α2-9 sialic acids (E-S001; QABio, Palm Desert, CA). After digestion, samples were separated from the sialidase by binding onto a LudgerClean Post-Exoglycosidase clean-up plate (Ludger Ltd.) for 60 min followed by elution of the glycans from the plate with water. The samples were analysed by HILIC-HPLC.

## HPLC analysis

HILIC-HPLC of 2-AB labelled glycans was performed using the 4.6 x 150 mm LudgerSep-N2 column (Ludger Ltd) on a 2795 HPLC with a 2475 fluorescence detector, controlled by Empower software (Waters, Manchester, UK). The flow rate was 1 mL/min. Solvent A was 50 mM ammonium formate pH 4.4, solvent B was acetonitrile. Samples were injected in 20% water / 80% acetonitrile; injection volume 25 µL. Gradient conditions were 20% - 38% solvent A at flow rate of 1.0 mL/min over 36 min, followed by 2 min at 100% solvent A, then finishing with 20% solvent A, giving a total run time of 45 min.

Waters GPC software with a cubic spline fit was used to allocate GU values to peaks and 2-AB labelled glucose homopolymer (CAB-GHP-30, Ludger Ltd) was used as a system suitability standard as well as an external calibration standard for GU allocation.[14]

## LC-ion trap-MS/MS analysis

Nano liquid chromatography (LC) was performed using an Ultimate 3000 LC system (Thermo Scientific; Breda, The Netherlands). Samples were applied to a C18 trapping column (C18 Acclaim PepMap100, 5 µm 0.1 mm × 20 mm; Thermo Scientific) and washed with 100% solvent A (0.1% formic acid in water) at a flow rate of 15 µL/min for 2 min prior to separation on a C18-RP analytical column (C18 Acclaim PepMapRSLC 100 Å, 2 µm, 75 µm x 150mm; Thermo Scientific) at a flow rate of 500 nL/min and elution was monitored by UV absorption (215 nm). Separation was achieved with 5 min 0% of solvent B (95% ACN, 5% water) as starting condition and a gradient to 25% in 15 min and in further 5 min to 70% of solvent B.

Afterwards, the column was washed for 5 min with 70% solvent B. The LC system was coupled via an online ESI-nanosprayer source to an Esquire AmaZonSpeed

ETD ESI-ion trap-MS (Bruker Daltonics). Capillary tubing with an inner diameter of 20 µm was used for nano-electrospray (ca 1200 V) and the solvent was evaporated at 250°C employing a nitrogen stream of 10 L/min. Spectra were acquired in positive-ion mode over a *m/z* range from 300 to 1800.

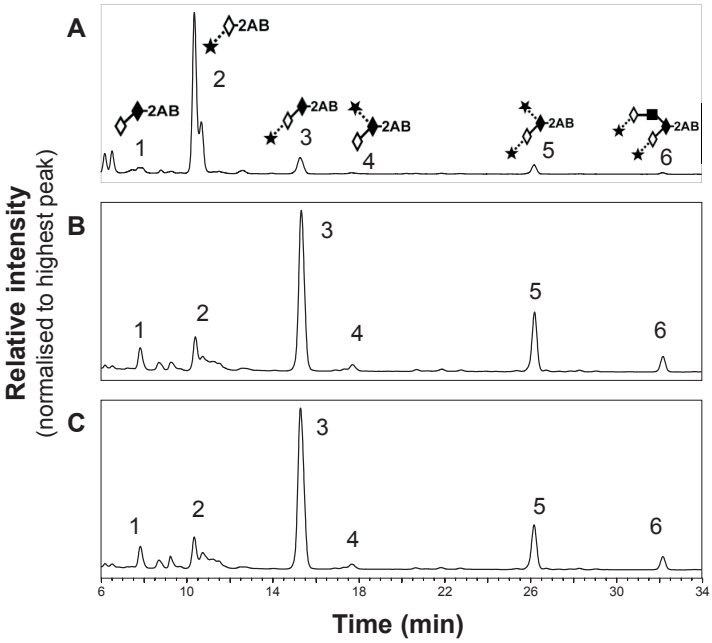



**Figure 1.** Comparison of HPLC profiles of 2-AB labelled fetuin O-glycans. The O-glycans were released with (A) hydrazine; (B) 50 mM EDTA salt in hydrazine; (C) 20 mM anhydrous EDTA in hydrazine. The following symbols are used to depict glycan structures[38]: ◊, galactose; ♦, *N*-acetylgalactosamine; ◈, fucose; ■, *N*-acetylglucosamine; ★, *N*-acetylneuraminic acid; ☆, *N*-glycolylneuraminic acid; dashed line, α-linkage; solid line, β-linkage.

| Structure | | Fetuin samples released with: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Hydrazine (n=6) | Hydrazine washed with 0.1% TFA prior to release (n=18) | Hydrazine washed with 100 mM EDTA prior to release (n=12) | 50 mM EDTA salt dihydrate in hydrazine (n=12) | 50 mM EDTA anhydrous in hydrazine (n=9) | 20 mM EDTA salt dihydrate in hydrazine (n=9) | 20 mM EDTA anhydrous in hydrazine (n=12) | 10 mM EDTA salt dihydrate in hydrazine (n=6) | 10 mM EDTA anhydrous in hydrazine (n=6) |
| 2AB | Average % area | **3.8** | **9.2** | **7.1** | **7.1** | **7.7** | **6.1** | **7.9** | **5.4** | **5.8** |
| | *STDEV* | *2.4* | *2.2* | *1.2* | *1.7* | *1.8* | *1.3* | *1.9* | *0.6* | *1.0* |
| | p-value | - | 0.36 | 0.35 | 0.09 | 0.06 | 0.03 | 0.09 | 0.07 | 0.13 |
| 2AB | Average % area | **71.3** | **19.1** | **17.4** | **24.5** | **20.7** | **28.3** | **21.8** | **26.2** | **23.6** |
| | *STDEV* | *23.3* | *7.3* | *3.0* | *4.6* | *3.1* | *3.3* | *2.9* | *2.3* | *1.2* |
| | p-value | - | **0.0** | **0.01** | 0.03 | 0.01 | 0.00 | 0.02 | 0.05 | 0.03 |
| 2AB | Average % area | **16.1** | **49.8** | **51.8** | **48.3** | **51.7** | **46.6** | **53.1** | **49.0** | **53.7** |
| | *STDEV* | *15.5* | *4.3* | *2.9* | *4.0* | *3.0* | *3.0* | *4.8* | *1.4* | *1.3* |
| | p-value | - | **0.01** | **0.01** | 0.03 | 0.01 | 0.00 | 0.01 | 0.04 | 0.01 |
| 2AB | Average % area | **1.1** | **2.6** | **2.7** | **2.3** | **2.7** | **2.6** | **2.4** | **2.7** | **2.4** |
| | *STDEV* | *0.7* | *1.0* | *0.4* | *0.7* | *0.5* | *0.6* | *0.6* | *0.2* | *0.2* |
| | p-value | - | 0.06 | **0.05** | 0.18 | 0.02 | 0.00 | 0.10 | 0.00 | 0.02 |
| 2AB | Average % area | **6.5** | **14.1** | **16.8** | **14.2** | **13.5** | **13.1** | **11.6** | **13.4** | **11.3** |
| | *STDEV* | *3.8* | *2.6* | *1.7* | *2.4* | *1.0* | *1.2* | *1.0* | *1.3* | *1.6* |
| | p-value | - | **0.00** | **0.01** | 0.05 | 0.03 | 0.03 | 0.09 | 0.14 | 0.25 |
| 2AB | Average % area | **1.2** | **3.6** | **4.2** | **3.5** | **3.6** | **3.2** | **3.1** | **3.4** | **3.1** |
| | *STDEV* | *1.1* | *1.0* | *0.6* | *0.7* | *0.4* | *0.4* | *0.3* | *0.4* | *0.4* |
| | p-value | - | **0.01** | **0.01** | 0.05 | 0.02 | 0.02 | 0.04 | 0.09 | 0.13 |

**Table 1.** Evaluation of EDTA addition during hydrazinolysis for O-glycosylation analysis of fetuin. Fetuin O-glycans were released by hydrazinolysis after pre-treatments (TFA wash or EDTA wash) or with addition of EDTA. O-glycan profiles were analysed by HILIC-UPLC of 2-AB labelled glycans. Relative abundances and standard deviations are given. Relative abundances of samples with TFA or EDTA treatment were compared to non-treated samples by t-test, and the resulting p-values are listed. Changes were significant for the p-value ≤ 0.05. Data for O-glycans released from fetuin that had been buffer exchanged with 0.1% TFA or 100 mM EDTA were taken from[35].

# Results and discussion

## HPLC analysis of 2-AB labelled O-glycans

We used bovine fetuin as a model to optimize the release and minimize the formation of peeling product. The O-glycans from dried bovine fetuin were released using a range of solutions: anhydrous hydrazine; 50 mM EDTA salt in hydrazine; 50 mM anhydrous EDTA in hydrazine; 20 mM EDTA salt in hydrazine; 20 mM anhydrous EDTA in hydrazine; 10 mM EDTA salt in hydrazine and 10 mM anhydrous EDTA in hydrazine. Released O-glycans were 2-AB labelled and analysed by HILIC-HPLC. For comparison two fetuin glycoprotein samples were buffer exchanged with 0.1% TFA or 100 mM EDTA prior to release with anhydrous hydrazine.[35] The HPLC data for this experiment is shown in Figure. 1 and Supplemental Figure. 1. The O-glycan structures from fetuin were consistent with published data.[36] The fetuin profile contains: core 1 Galα1-3GalNAc (peak 1), mono-sialylated and di-sialylated core 1 O-

glycans: Neu5Acα2-3Galβ1-3GalNAc (peak 3); Neu5Acα2-6(Galβ1-3)GalNAc (peak 4); di-sialylated core 1 O-glycan Neu5Acα2-3Galβ1-3(Neu5Acα2-6)GalNAc (peak 5); di-sialylated core 2 structure Neu5Acα2-3Galβ1-3(Neu5Acα2-3Galβ1-4GlcNAcβ1-6)GalNAc (peak 6) as well as the peeled product Neu5Acα2-3Gal (peak 2) (Supplemental Table 1).

The highest amount of the peeled product (71%) was observed for the fetuin sample that was incubated with anhydrous hydrazine only, without prior washes (peak 2, Figure. 1A). For the samples washed with 0.1%TFA or 100 mM EDTA, the degree of peeling was consistent with published results [35] with an average relative abundance of 19% and 17%, respectively (peak 2, Supplemental Figure. 1B, C, Table 1.). The amount of peeling for fetuin samples that were released with a range of low-molarity solutions of EDTA salt or anhydrous EDTA in hydrazine was between 20% and 28% (Table 1.). The most efficient reduction of peeling was observed for the samples released with 50 mM and 20 mM EDTA anhydrous in hydrazine (22% and 21% relative intensity for peak 2; Supplemental Figure. 1E, G; Table 1.).

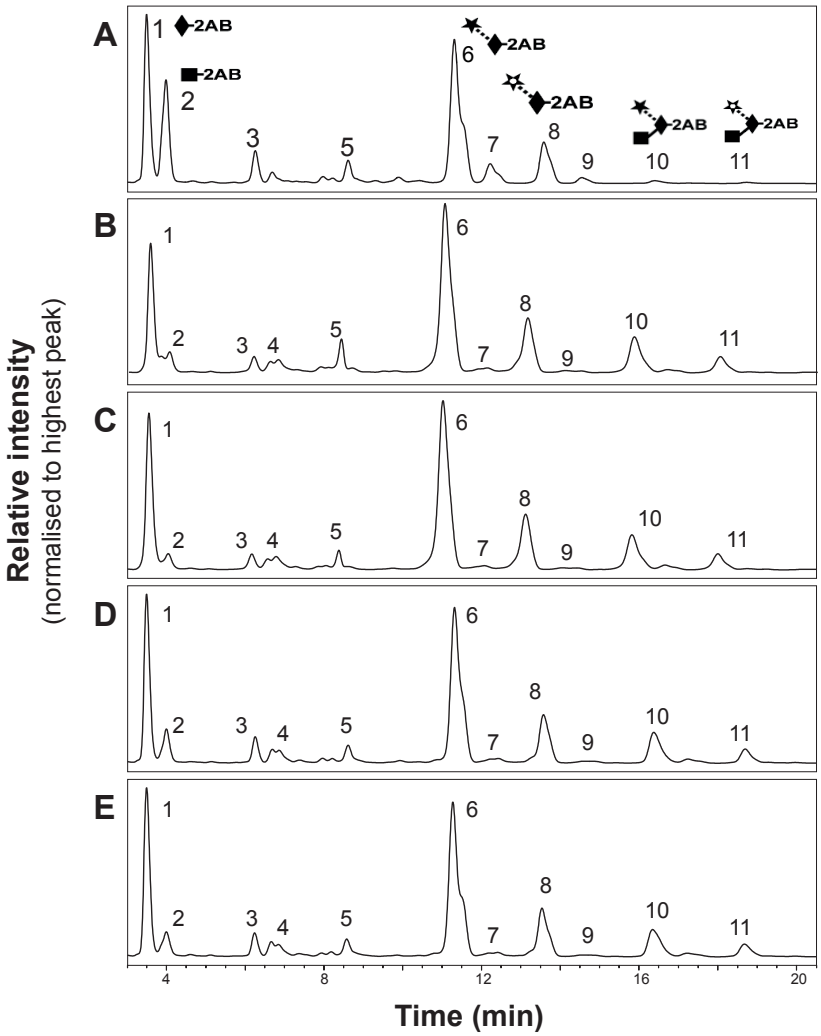



**Figure 2.** Comparison of HPLC profiles of 2-AB labelled BSM O-glycans. The O-glycans were released with (A) hydrazine; (B) sample was buffer exchange with 0.1% TFA prior to hydrazinolysis; (C) sample was buffer exchange with 100 mM EDTA prior to hydrazinolysis; hydrazinolysis was performed in (D) 50 mM EDTA salt in hydrazine or (E) 20 mM anhydrous EDTA in hydrazine.

| Structure | | BSM samples released with: | | | | |
|---|---|---|---|---|---|---|
| | | Hydrazine (n=6) | Hydrazine | | 50 mM EDTA salt dihydrate in hydrazine (n=6) | 20 mM EDTA anhydrous in hydrazine (n=6) |
| | | | washed with 0.1% TFA prior to release (n=9) | washed with 100 mM EDTA prior to release (n=9) | | |
| | **Average % area** | **22.3** | **18.2** | **18.8** | **22.0** | **24.0** |
| | *STDEV* | - | *1.4* | *2.2* | *0.7* | *2.4* |
| | p-value | - | 0.84 | 0.51 | 0.50 | 0.35 |
| | **Average % area** | **16.6** | **4.8** | **4.2** | **5.7** | **5.7** |
| | *STDEV* | *0.5* | *1.3* | *0.9* | *0.5* | *1.8* |
| | p-value | - | **0.01** | **0.00** | **0.00** | **0.01** |
| **fucosylated structure** | **Average % area** | **4.7** | **0.2** | **3.5** | **3.7** | **4.0** |
| | *STDEV* | - | *0.2* | *0.6* | *0.0* | *0.8* |
| | p-value | - | **0.05** | 0.07 | **0.00** | 0.26 |
| | **Average % area** | **-** | **2.4** | **2.3** | **2.1** | **1.5** |
| | *STDEV* | - | *0.7* | *0.2* | *0.1* | *0.8* |
| | p-value | - | **0.00** | **0.00** | **0.00** | 0.08 |
| **sialylated structure** | **Average % area** | **4.0** | **4.3** | **2.9** | **3.4** | **3.4** |
| | *STDEV* | - | *0.5* | *0.3* | *0.1* | *0.4* |
| | p-value | - | 0.38 | **0.01** | **0.01** | 0.08 |
| | **Average % area** | **34.8** | **39.5** | **41.0** | **36.9** | **36.5** |
| | *STDEV* | - | *0.7* | *1.2* | *0.5* | *2.0* |
| | p-value | - | **0.00** | **0.00** | **0.02** | 0.27 |
| **sialylated structure** | **Average % area** | **5.0** | **1.5** | **1.6** | **2.0** | **1.9** |
| | *STDEV* | - | *0.5* | *0.4* | *0.3* | *0.1* |
| | p-value | - | **0.00** | **0.00** | **0.00** | **0.00** |
| | **Average % area** | **9.9** | **13.1** | **13.3** | **11.8** | **11.6** |
| | *STDEV* | - | *0.3* | *0.4* | *0.2* | *0.9* |
| | p-value | - | **0.00** | **0.00** | **0.00** | 0.08 |
| **sialylated structure** | **Average % area** | **1.4** | **0.8** | **0.7** | **0.9** | **0.8** |
| | *STDEV* | - | *0.3* | *0.2* | *0.0* | *0.1* |
| | p-value | - | **0.01** | **0.01** | **0.00** | **0.00** |
| | **Average % area** | **0.8** | **8.3** | **8.3** | **8.0** | **7.4** |
| | *STDEV* | - | *0.9* | *0.9* | *0.5* | *1.0* |
| | p-value | - | **0.00** | **0.00** | **0.00** | **0.01** |
| | **Average % area** | **0.3** | **3.5** | **3.5** | **3.4** | **3.2** |
| | *STDEV* | *0.0* | *0.4* | *0.4* | *0.2* | *0.5* |
| | p-value | - | **0.00** | **0.00** | **0.00** | **0.01** |

**Table 2.** Evaluation of EDTA addition during hydrazinolysis for O-glycosylation analysis of bovine submaxillary mucin (BSM). BSM O-glycans were released by hydrazinolysis after pre-treatments (TFA wash or EDTA wash) or with addition of EDTA. O-glycan profiles were analysed by HILIC-UPLC of 2-AB labelled glycans. Relative abundances and standard deviations are given. Relative abundances of samples with TFA or EDTA treatment were compared to non-treated samples by t-test, and the resulting p-values are listed. Changes were significant for the p-value ≤ 0.05. Data for O-glycans released from BSM exchanged with 0.1% TFA or 100 mM EDTA were taken from[35].

Based on the results for fetuin we analysed the O-glycosylation of a second model glycoprotein, BSM.[39] Mucin O-glycans were released with anhydrous hydrazine, 50 mM EDTA salt and 20 mM anhydrous EDTA solution in hydrazine, 2-AB labelled and analysed by HILIC-HPLC.

The HPLC results were compared with previously published data for the BSM samples cleaned up prior to hydrazinolysis with 0.1% TFA and 100 mM EDTA [35]. The identities of the observed peaks were confirmed by comparison with previously reported data.[36,40,41] Six different glycan species were identified: GalNAc (peak 1), GlcNAcβ1-3GalNAc (peak 3), Neu5Acα2-6GalNAc (peak 6); Neu5Gcα2-6GalNAc (peak 8); GlcNAcβ1-3(Neu5Acα2-6)GalNAc (peak 10) and GlcNAcβ1-3(Neu5Gcα2-6)GalNAc (peak 11) along with peeled product GlcNAc (peak 2) (Supplemental Table 1). Moreover, BSM was also found to contain additional sialylated species (peaks 5, 7 and 9) and a fucosylated structure (peak 3) which were not fully structurally characterised.

The highest relative abundance of unwanted peeled product was detected for the sample that was released with anhydrous hydrazine (17%, peak 2, Figure. 2A; Table 2). For the samples released with solutions of 50 mM EDTA salt and 20 mM anhydrous EDTA in hydrazine the degree of degradation product was much lower (6%, peak 2, Figure. 2D, E; Table 2) and consistent with published results.[35]

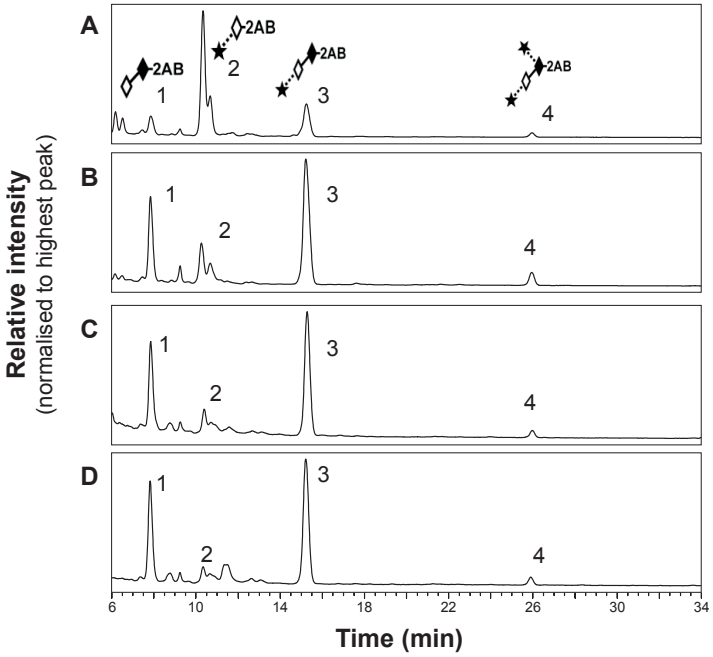



**Figure 3.** Comparison of HPLC profiles of 2-AB labelled serum IgA O-glycans. The O-glycans were released with (A) hydrazine; (B) sample was buffer exchange with 0.1% TFA prior to hydrazinolysis; release was performed with (C) 50 mM EDTA salt in hydrazine and (D) 20 mM anhydrous EDTA in hydrazine.

To further evaluate the improved hydrazinolysis protocol, more experiments were performed using IgA purified from human plasma. The O-glycans were released from IgA samples using the following conditions: (i) sample without any clean-up was released with anhydrous hydrazine; (ii) sample was cleaned by centrifugal filtration with 0.1% TFA before release with anhydrous hydrazine according to previously described procedure [35]; (iii) sample without any clean-up was released with 50 mM EDTA salt in hydrazine; (iv) sample without any clean-up was released with 20 mM anhydrous EDTA in hydrazine. The HPLC chromatograms obtained from these four IgA experiments are shown in Figure. 3. The IgA O-glycan profiles contain four major peaks corresponding to: core 1 Galα1-3GalNAc (peak 1), mono-sialylated and di-sialylated core 1 O-glycans: Neu5Acα2-3Galβ1-3GalNAc (peak 3); di-sialylated core

1 O-glycan Neu5Acα2-3Galβ1-3(Neu5Acα2-6)GalNAc (peak 4), and the sialylated peeled product Neu5Acα2-3Gal (peak 2, Figure. 3.) (Supplemental Table 1). The mono-sialylated and di-sialylated core 1 (peaks 3 & 4, Figure. 4) digested to the neutral core 1 (peak 1, Figure. 4) and the sialylated peeled product peak was digested to galactose (peak 2, Figure. 4A) following treatment with sialidase. These structures are in agreement with previously published structures.[37,42]

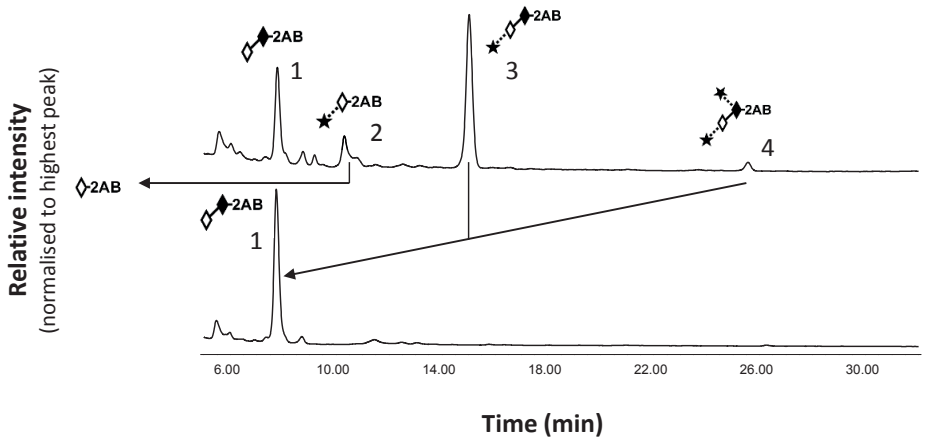



**Figure 4.** Glycan sequencing of IgA O-glycans before and after treatment with sialidase. The O-glycans were released with 20 mM anhydrous EDTA in hydrazine and 2-AB labelled. Aliquots of the total O-glycan pool were incubated with sialidase. O-glycans were analyzed by HILIC-HPLC before (A) and after sialidase treatment. Arrows indicate digestion pathways.

A significant difference in the relative abundance of peeling products was apparent between the different reagents used for the release. The highest occurrence of peeling was observed for the samples released with anhydrous hydrazine that did not have any EDTA added (68%, peek 2, Figure. 3A, Table 3).

| Structure | | IgA samples released with: | | | |
|---|---|---|---|---|---|
| | | Hydrazine (n=3) | Hydrazine washed with 0.1% TFA prior to release (n=3) | 50 mM EDTA salt dihydrate in hydrazine (n=6) | 20 mM EDTA anhydrous in hydrazine (n=6) |
| | **Average % area** | **8.2** | **21.5** | **29.8** | **31.5** |
| | *STDEV* | - | *2.4* | *0.9* | *0.3* |
| | p-value | - | 0.08 | 0.00 | 0.01 |
| | **Average % area** | **68.1** | **25.5** | **14.1** | **12.2** |
| | *STDEV* | - | *5.9* | *0.7* | *2.4* |
| | p-value | - | 0.06 | 0.00 | 0.02 |
| | **Average % area** | **21.5** | **48.9** | **53.0** | **53.4** |
| | *STDEV* | - | *2.8* | *0.3* | *2.0* |
| | p-value | - | 0.05 | 0.00 | 0.03 |
| | **Average % area** | **2.3** | **4.2** | **3.1** | **3.0** |
| | *STDEV* | - | *0.7* | *0.1* | *0.1* |
| | p-value | - | 0.16 | 0.01 | 0.09 |

**Table 3.** Evaluation of EDTA addition during hydrazinolysis for O-glycosylation analysis of serum IgA. IgA O-glycans were released by hydrazinolysis after TFA wash or with addition of EDTA. O-glycan profiles were analysed by HILIC-UPLC of 2-AB labelled glycans. Relative abundances and standard deviations are given. Relative abundances of samples with TFA or EDTA treatment were compared to non-treated samples by t-test, and the resulting p-values are listed. Changes were significant for the p-value ≤ 0.05.

| Structure | Composition | | | Mass calculated $[M+H]^+$ | Mass registered $[M+H]^+$ | | |
|---|---|---|---|---|---|---|---|
| | | | | | IgA samples released with : | | |
| | Hex | HexNAc | Neu5Ac | | anhydrous hydrazine | 50mM EDTA (salt) in hydrazine | 20mM EDTA (anhydrous) in hydrazine |
| | 1 | 1 | 0 | 504.22 | nd | 504.26 | 504.27 |
| | 1 | 0 | 1 | 592.23 | 592.27 | nd | nd |
| | 1 | 1 | 1 | 795.31 | 795.37 | 795.36 | 795.35 |

**Table 4.** Mass spectrometric data and corresponding structures of the IgA O-glycans peaks. nd, not detected.

## LC-ion trap-MS/MS analysis of 2-AB labelled IgA samples.

To confirm the identities and the amounts of observed peaks (peaks 1, 2 and 3, Figure. 3) the 2-AB labelled O-glycans released from serum IgA were fractionated by HILIC-HPLC and characterized by LC-ion trap-MS/MS.

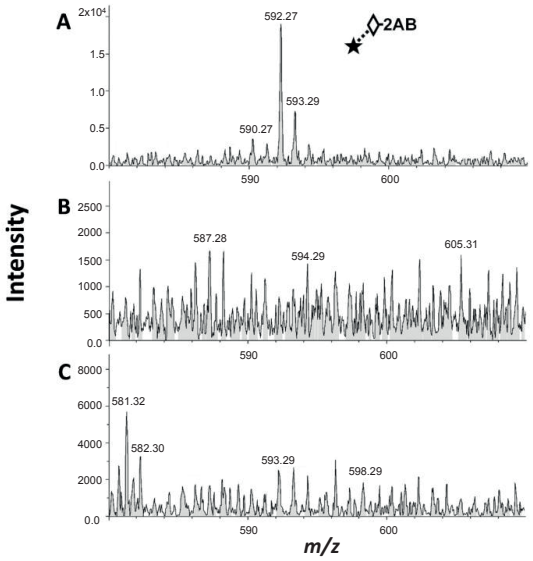



**Figure 5.** LC-ESI-MS mass spectra of collected peak 2 (peeled product) of the three IgA O-glycan samples generated with (A) hydrazine; (B) 50 mM EDTA salt in hydrazine; (C) 20 mM anhydrous EDTA in hydrazine (see Fig. 4). Spectra were summed over the entire elution range of the peeling product.

The observed molecular masses and the proposed structures based on previously reported data [37,42] [40-41] are summarized in Table 4. The results from LC-MS showed high level of peeled product for the samples released with anhydrous hydrazine (Figure 5A). For the IgA samples released with solutions of 50 mM EDTA salt and 20 mM anhydrous EDTA in hydrazine the peeled product was not detected (Figure 5B & C).

## Conclusions

Here we present an improved hydrazinolysis protocol for the release of O-glycans (in their non-reduced form) from glycoproteins. We demonstrate that undesirable peeling can be efficiently reduced by adding EDTA directly to the hydrazine. These results are comparable with previously published data where the

samples are cleaned up with low-molarity EDTA solutions prior to hydrazine release [35] suggesting that the chelation of cations may be beneficial for suppressing O-glycan peeling.

We applied the current method to the analysis of O-glycans from bovine fetuin, bovine submaxillary gland mucin (BSM) and serum IgA and found that undesirable peeling was reduced from 71% for the fetuin O-glycans released with anhydrous hydrazine to 17% for the same sample released with solutions of 20 mM anhydrous EDTA or 50 mM EDTA salt in hydrazine. Significant reductions in the amount of peeled product were also observed for the BSM and human serum IgA samples. For the BSM, the peeled product decreased from 17% to 6% and for the IgA sample it decreased from 68% to 12%.

When compared to the previously published protocol [35] which relied on several consecutive washes prior to O-glycan release, this improved method has a reduced number of sample handling steps which make the process much quicker and helps prevent sample losses when working with low amounts of materials.

## Acknowledgements

The authors thank Carolien A. Koeleman and Agnes L. Hipgrave-Ederveen for LC-ion trap-MS/MS analysis. This work was supported by the European Union's Seventh Framework Program (FP7-Health-F5-2011) under grant agreement n°278535 (HighGlycan).

## Reference List:


(1) Houde, D.; Peng, Y.; Berkowitz, S. A.; Engen, J. R. *Molecular & cellular proteomics : MCP* **2010**, *9*, 1716-1728.

(2) Parodi, A. J. *Biochim Biophys Acta* **1999**, *1426*, 287-295.

(3) Varki, A. *Glycobiology* **1993**, *3*, 97-130.

(4) Brockhausen, I. *Biochim. Biophys. Acta* **1999**, *1473*, 67-95.

(5) Brockhausen, I.; Kuhns, W. *Glycoproteins and Human Disease*; CRC Press and Mosby Year Book, Chapman and Hall, New York, 1997.

(6) Brockhausen, I. *EMBO Rep* **2006**, *7*, 599-604.

(7) Chandrasekaran, E. V.; Xue, J.; Xia, J.; Locke, R. D.; Patil, S. A.; Neelamegham, S.; Matta, K. L. *Journal of proteome research* **2012**, *11*, 2609-2618.

(8) Tarelli, E.; Smith, A. C.; Hendry, B. M.; Challacombe, S. J.; Pouria, S. *Carbohydr Res* **2004**, *339*, 2329-2335.

(9) Julian, B. A.; Novak, J. *Current opinion in nephrology and hypertension* **2004**, *13*, 171-179.

(10) Berger, E. G. *Biochim Biophys Acta* **1999**, *1455*, 255-268.

(11) Zhang, P.; Wang, T.; Bardor, M.; Song, Z. *Pharmaceutical Bioprocessing* **2013**, *1*, 89-104.

(12) Shahrokh, Z.; Royle, L.; Saldova, R.; Bones, J.; Abrahams, J. L.; Artemenko, N. V.; Flatman, S.; Davies, M.; Baycroft, A.; Sehgal, S.; Heartlein, M. W.; Harvey, D. J.; Rudd, P. M. *Molecular pharmaceutics* **2011**, *8*, 286-296.

(13) Schiestl, M.; Stangler, T.; Torella, C.; Cepeljnik, T.; Toll, H.; Grau, R. *Nat. Biotechnol* **2011**, *29*, 310-312.

(14) Endo, T. *Biol. Pharm. Bull* **2009**, *32*, 765-766.

(15) Zauner, G.; Koeleman, C. A.; Deelder, A. M.; Wuhrer, M. *Journal of separation science* **2010**, *33*, 903-910.

(16) Iwase, H.; Hotta, K. *Methods Mol. Biol* **1993**, *14*, 151-159.

(17) Geyer, H.; Geyer, R. *Biochim. Biophys. Acta* **2006**, *1764*, 1853-1869.

(18) Bhavanandan, V. P.; Umemoto, J.; Davidson, E. A. *Biochemical and biophysical research communications* **1976**, *70*, 738-745.

(19) Iyer, R. N.; Carlson, D. M. *Arch. Biochem. Biophys* **1971**, *142*, 101-105.

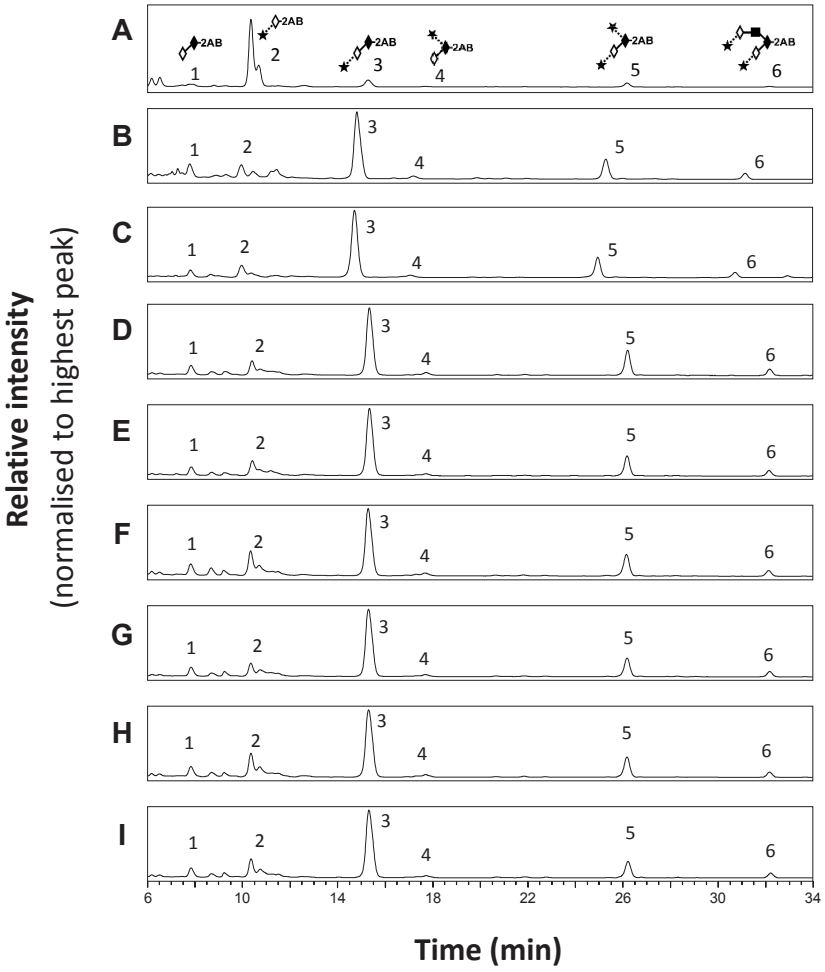
(20) Cooper, C. A.; Packer, N. H.; Redmond, J. W. *Glycoconj. J* **1994**, *11*, 163-167.

(21) Carlson, D. M. *J. Biol. Chem* **1968**, *243*, 616-626.

(22) Stadheim, T. A.; Li, H.; Kett, W.; Burnina, I. N.; Gerngross, T. U. *Nat. Protoc* **2008**, *3*, 1026-1031.
(23) Zaia, J. *OMICS* **2010**, *14*, 401-418.
(24) Koles, K.; van Berkel, P. H.; Pieper, F. R.; Nuijens, J. H.; Mannesse, M. L.; Vliegenthart, J. F.; Kamerling, J. P. *Glycobiology* **2004**, *14*, 51-64.
(25) Karlsson, N. G.; Schulz, B. L.; Packer, N. H. *J. Am. Soc. Mass Spectrom* **2004**, *15*, 659-672.
(26) Cooke, C. L.; An, H. J.; Kim, J.; Solnick, J. V.; Lebrilla, C. B. *Anal. Chem* **2007**, *79*, 8090-8097.
(27) Chai, W.; Feizi, T.; Yuen, C. T.; Lawson, A. M. *Glycobiology* **1997**, *7*, 861-872.
(28) Huang, Y.; Mechref, Y.; Novotny, M. V. *Anal. Chem* **2001**, *73*, 6063-6069.
(29) Yamada, K.; Hyodo, S.; Kinoshita, M.; Hayakawa, T.; Kakehi, K. *Anal. Chem* **2010**, *82*, 7436-7443.
(30) Turyan, I.; Hronowski, X.; Sosic, Z.; Lyubarskaya, Y. *Analytical biochemistry* **2014**, *446*, 28-36.
(31) Patel, T.; Bruce, J.; Merry, A.; Bigge, C.; Wormald, M.; Jaques, A.; Parekh, R. *Biochemistry* **1993**, *32*, 679-693.
(32) Merry, A. H.; Neville, D. C.; Royle, L.; Matthews, B.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M. *Anal. Biochem* **2002**, *304*, 91-99.
(33) Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M. *Anal. Bioanal. Chem* **2010**, *397*, 3457-3481.
(34) Liu, L.; Tharmalingam, T.; Maischberger, E.; Albrecht, S.; Gallagher, M. E.; Miranda-Casoluengo, R.; Meijer, W. G.; Rudd, P. M.; Irwin, J. A. *Glycoconjugate journal* **2013**, *30*, 791-800.
(35) Kozak, R. P.; Royle, L.; Gardner, R. A.; Fernandes, D. L.; Wuhrer, M. *Anal. Biochem* **2012**, *423*, 119-128.
(36) Bigge, J. C.; Patel, T. P.; Bruce, J. A.; Goulding, P. N.; Charles, S. M.; Parekh, R. B. *Anal. Biochem* **1995**, *230*, 229-238.
(37) Royle, L.; Mattu, T. S.; Hart, E.; Langridge, J. I.; Merry, A. H.; Murphy, N.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M. *Anal. Biochem* **2002**, *304*, 70-90.
(38) Harvey, D. J.; Merry, A. H.; Royle, L.; Campbell, M. P.; Rudd, P. M. *Proteomics* **2011**, *11*, 4291-4295.
(39) Spiro, R. G.; Bhoyroo, V. D. *J. Biol. Chem* **1974**, *249*, 5704-5717.
(40) Zauner, G.; Koeleman, C. A.; Deelder, A. M.; Wuhrer, M. *Biochim. Biophys. Acta* **2012**, *1820*, 1420-1428.

(41) Robbe, C.; Capon, C.; Flahaut, C.; Michalski, J. C. *Electrophoresis* **2003**, *24*, 611-621.
(42) Mattu, T. S.; Pleass, R. J.; Willis, A. C.; Kilian, M.; Wormald, M. R.; Lellouch, A. C.; Rudd, P. M.; Woof, J. M.; Dwek, R. A. *The Journal of biological chemistry* **1998**, *273*, 2260-2272.

# Supplementary information




**Supplementary Figure 1.** Fetuin O-glycan profiling after hydrazinolysis. (A) Sample was subjected to conventional hydrazinolysis without pretreatment; (B) sample was buffer exchanged with 0.1% TFA prior to hydrazinolysis; (C) sample was buffer exchanged with 100mM EDTA prior to hydrazinolysis. In (D-I) the following additives were used with hydrolysis: (D) 50mM EDTA (salt) in hydrazine; (E) 50mM EDTA (anhydrous) in; (F) 20mM EDTA (salt) in hydrazine; (G) 20mM EDTA (anhydrous) in hydrazine; (H) 10mM EDTA (salt) in hydrazine; (I) 10mM EDTA (anhydrous) in hydrazine. The following symbols are used to depict glycan structures[37]: ◊, galactose; ♦, *N*-acetylgalactosamine; ◈, fucose; ■, *N*-acetylglucosamine; ★, *N*-acetylneuraminic acid; ☆, *N*-glycolylneuraminic acid; dashed line, α-linkage; solid line, β-linkage.

# Chapter 5

## Comparison of procainamide and 2-aminobenzamide labelling for profiling and identification of glycans by LC-FLR-ESI-MS.

**Radoslaw P. Kozak[1], Concepcion Badia Tortosa[1], Daryl L. Fernandes[1] and Daniel I. R. Spencer[1]**

[1]Ludger Ltd., Culham Science Centre, Oxfordshire, OX14 3EB, United Kingdom

Analytical Biochemistry (2015) 38-40.

## Abstract

One of the most widely used methods for glycan analysis is fluorescent labelling of released glycans followed by HILIC-(U)HPLC. Here, we compare the data obtained by (U)HPLC-FLR coupled to ESI-MS for procainamide and 2-AB labelled N-glycans released from human IgG. Fluorescence profiles from procainamide show comparable chromatographic separation to those obtained for 2-AB but gave higher fluorescence intensity as well as significantly improved ESI ionization efficiency (up to 30 times that of 2-AB). Thus, labelling with procainamide increases the ability to identify minor glycan species which may have significant biological activity.

## Introduction

Protein glycosylation is a post-translational modification that influences many protein functions.[1,2] Detailed monitoring and controlling glycosylation is essential in biopharmaceutical development and quality control of drugs that are glycosylated (e.g., erythropoietin, EPO).[3,4]

A number of analytical techniques have been developed for glycan characterization including high-performance liquid chromatography (HPLC), mass spectrometry (MS) and high pH anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD). Each of these techniques can provide different information and a combination of these techniques in an LC-MS system can greatly improve full glycan characterization.[5]

HPLC analysis of released and fluorescently labelled (via reductive amination) N-glycans is one of the most common methods; the stoichiometric attachment of one label per glycan allows the relative quantitation of different glycan species based on fluorescence or UV-absorbance intensity.[6] This approach, combined with specific enzyme digestions can help to assign glycan structures more fully.

MS and MS/MS can provide composition and sequence data, however samples containing isomeric and isobaric glycan structures might not be distinguished by this technique.

HPLC-FLR coupled to ESI-MS (electrospray ionization) system can greatly improve glycan analysis by providing specific structural information.

There are several fluorescent labels that have been used for the reductive amination of N-glycans, such as 2-aminobenzamide (2-AB), 2-aminobenzoic acid (2-AA), 1-aminopyrene-3,6,8-trisulfonic acid (APTS), 2-aminopyridine (AP), procaine or

procainamide.[7] 2-AB is the most widely used label providing relative quantitation of N-glycan structures through efficient labelling and the production of highly stable labelled glycans. However, a major drawback of 2-AB is poor ionization efficiency when analyzing N-glycan structures by ESI-MS.[5]

Yoshino et al.[8] and Takao et al.[9] reported a 1000-fold increase in sensitivity over free oligosaccharides for glycans labelled with procaine (2-(diethylamino)ethyl 4-aminobenzoate) in matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) and ESI-MS. However, Pabst et al.[7] found that the performance of procaine in NP–HPLC was disappointing. The considerable differences between elution times and separation of glycans labelled with procaine and those labelled with 2-AB were observed.

Here we present data comparing two fluorescent labels for quantitative N-glycan profiling and identification by (U)HPLC-FLR coupled to ESI-MS and MS/MS. N-glycans from human polyclonal IgG were released, 2-AB labelled or procainamide (4-amino-N-(2-diethylaminoethyl) benzamide) labelled and purified on a liquid-handling robotic platform.

## Materials and Methods

### Materials

All reagents and kits for release, N-glycan labelling and clean-up were from Ludger Ltd. (Oxford, UK). Acetonitrile (Romil; 190 SpS for UV/ gradient quality) was obtained from Charlton Scientific (Charlton, Oxon, UK). Human IgG was from Ludger

Ltd (Oxford, UK) and all other reagents were obtained from Sigma-Aldrich (Dorset, UK).

Glycan release, 2-AB labelling, procainamide labelling and labelled glycan purification were performed on a Hamilton Microlab STARLet liquid-handling robotic platform.[10]

## Release of N-glycans

Samples were denatured by heating for 7 min at 99°C in 2% SDD/1M β-mercaptoethanol, before overnight incubation at 37°C with PNGase F (Ludger Ltd.) in the presence of Triton X-100. The released N-glycans were converted to aldoses with 0.1% formic acid,[11] filtered through a protein binding plate (Ludger Ltd.) and dried.

## Fluorescent labelling

Released N-glycans were fluorescently labelled by reductive amination in 10 µL of water with either 10 µL of 2-AB labelling solution as per the Ludger 2-AB glycan labelling kit containing 2-picoline borane or 10 µL of procainamide labelling solution as per Ludger procainamide glycan labelling kit containing 2-picoline borane. Samples were incubated at 65°C for 1 hour.

## Purification of 2-AB labelled glycans

The 2-AB labelled N-glycans were cleaned up using Ludger Clean T1 Cartridges. 2-AB labelled N-glycans were eluted with water (1ml). The samples were dried by vacuum centrifugation and re-suspended in water (100 µL) for further analysis.

## Purification of Procainamide labelled glycans

The removal of unreacted procainamide dye was performed using a Ludger Clean plate for cleanup of procainamide labelled glycans. The procainamide labelled N-glycans were eluted with water (200 µL). The samples were dried by vacuum centrifugation and re-suspended in water (100 µL) for further analysis.

## LC-ESI-MS and MS/MS analysis

2-AB and procainamide labelled samples were analysed by LC-ESI-MS. 25 µL of each sample was injected onto an ACQUITY UPLC® BEH-Glycan 1.7 µm, 2.1 x 150 mm column at 40°C on the Dionex Ultimate 3000 UHPLC instrument with a fluorescence detector ($\lambda$ex = 250nm, $\lambda$em = 428nm for 2-AB and $\lambda$ex = 310nm, $\lambda$em = 370nm for procainamide), attached to a Bruker Amazon Speed ETD. The running conditions used were: Solvent A was 50 mM ammonium formate pH 4.4 made from Ludger Stock Buffer (Ludger), and solvent B was acetonitrile. Gradient conditions were: 0 to 38.5 min, 76% to 58 % B; 38.5 to 40.5 min, 58 to 40.5 % B at flow rate 0.4mL/min. 40.5 to 42.5 min, 40 % B at a flow rate of 0.25 mL/min; 42.5 to 44.5 min, 40 to 76 % B at a flow rate of 0.25 mL/min; 44.5 to 50.5 min, 76% B at a flow rate of 0.25 mL/min. 50.5 to 51.5 min, 76% B at a flow rate 0.25 mL/min; 51.5 to 55 min, 76% B at a flow rate 0.4 mL/min. The Amazon Speed settings used were: source temperature 250 °C, gas flow 10 L/min; Capillary voltage 4500 V; ICC target 200,000; max accu time 50.00 ms; rolling average 2; number of precursors ions selected 3, release after 0.2 min; Positive ion mode; Scan mode: enhanced resolution; mass range scanned, 200-1500; Target mass, 900. A glucose homopolymer ladder (Ludger), labelled with either 2AB or procainamide was used as

a system suitability standard as well as an external calibration standard for GU allocation.[11]

## Results and Discussion

Release and recovery of N-glycans with free reducing termini is important for structural and functional analysis. Using human IgG as a model glycoprotein, we compared two fluorescent labels, 2-aminobenzamide (2-AB) and procainamide. The N-glycans from human IgG were released, labelled and purified using a liquid-handling robotic platform.

The profiles obtained by (U)HPLC-FLR coupled to ESI-MS from the human IgG samples labelled with 2-AB and procainamide were compared in Figure 1. Comparable N-glycan profiles were obtained for each sample, labelled with 2-AB and procainamide (Figure 1a and 1d), as described before.[12] The peak intensity of the fluorescent profiles from the samples labelled with procainamide (the emission and excitation wavelengths were set to λex 310nm and λem 370nm) were higher than those from 2-AB labelled (λex 250nm and λem 428nm) allowing better integration of the lower intensity peaks.

The base peak chromatogram (BPC) in positive mode ESI-MS for procainamide labelled human IgG N-glycans showed up to 30 times higher signal intensity compared to 2-AB (Figure 1b and 1e). The extracted ion chromatogram (EIC) profiles for the 2-AB labelled (Figure 1c) and procainamide labelled (Figure 1f) A2BG2S2 glycan (Table 1, Figure 1a and 1d, peak 17) showed significantly higher ion intensity for procainamide derivative, enabling identification of low level glycan species not mass detected using 2-AB. This high sensitivity in positive mode ESI-MS

can be explained because of the high proton affinity of the procainamide basic tail (2-(diethylamino)ethyl group).

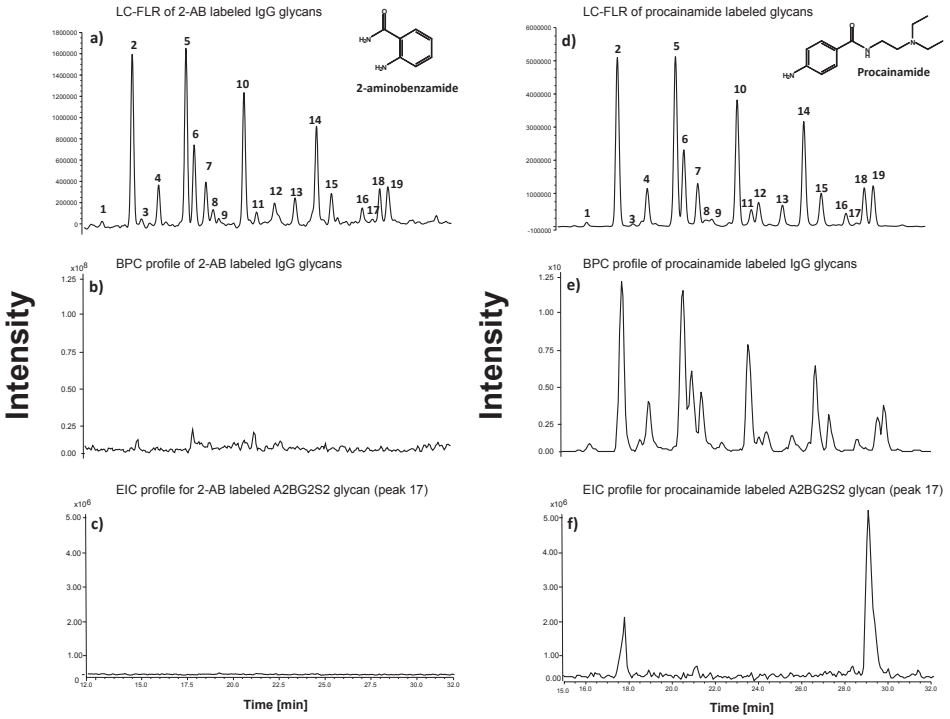



**Figure 1.** Comparison of fluorescently labelled human IgG glycans. (a) LC-FLR of 2-AB labelled IgG glycans, (b) LC-ESI-MS BPC profile of 2-AB labelled IgG glycans, (c) LC-ESI-MS EIC for 2-AB labelled A2BG2S2 glycan (peak 17 on LC-FLR trace), (d) LC-FLR of procainamide labelled glycans, (e) LC-ESI-MS BPC profile of procainamide labelled IgG glycans, (f) LC-ESI-MS EIC for procainamide labelled A2BG2S2 glycan (peak 17 on LC-FLR trace).

All of the 19 peaks that were detected by (U)HPLC-FLR were identified by MS and MS/MS (Table 1). One, two or three charge states were detected for the procainamide derivatives. In some cases, when the MS/MS analysis was performed on singly charged precursor ions, a loss of 73 Da was observed. This observation

indicated degradation of the procainamide label during MS/MS fragmentation and corresponds to loss of a diethylamine ion.

In summary, the comparison of the two fluorescent labels presented here demonstrates that procainamide derivatization can be used as a standard method for manual and automated glycan characterization. The procainamide labelled glycans are suitable for both (U)HPLC-FLR analysis, providing good chromatographic separation for relative quantitation and ESI-MS analysis, providing more efficient ionization for glycan identification.

## Acknowledgements

The authors thank Richard Gardner for setting up the high throughput release, labelling and clean-up methods on the Hamilton robot and Louise Royle for expert assistance and critically reading the manuscript.

This work was supported by the European Union's Seventh Framework Program (FP7-Health-F5-2011) under grant agreement n°278535 (HighGlycan).

## Reference List:


(1) Parodi, A. J. *Biochim Biophys Acta* **1999**, *1426*, 287-295.

(2) Houde, D.; Peng, Y.; Berkowitz, S. A.; Engen, J. R. *Molecular & cellular proteomics : MCP* **2010**, *9*, 1716-1728.

(3) Zhang, P.; Wang, T.; Bardor, M.; Song, Z. *Pharmaceutical Bioprocessing* **2013**, *1*, 89-104.

(4) Shahrokh, Z.; Royle, L.; Saldova, R.; Bones, J.; Abrahams, J. L.; Artemenko, N. V.; Flatman, S.; Davies, M.; Baycroft, A.; Sehgal, S.; Heartlein, M. W.; Harvey, D. J.; Rudd, P. M. *Molecular pharmaceutics* **2011**, *8*, 286-296.

(5) Harvey, D. J. *Journal of the American Society for Mass Spectrometry* **2000**, *11*, 900-915.

(6) Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M. *Anal. Bioanal. Chem* **2010**, *397*, 3457-3481.

(7) Pabst, M.; Kolarich, D.; Poltl, G.; Dalik, T.; Lubec, G.; Hofinger, A.; Altmann, F. *Anal. Biochem* **2009**, *384*, 263-273.

(8) Yoshino, K.; Takao, T.; Murata, H.; Shimonishi, Y. *Analytical chemistry* **1995**, *67*, 4028-4031.

(9) Takao, T.; Tambara, Y.; Nakamura, A.; Yoshino, K.; Fukuda, H.; Fukuda, M.; Shimonishi, Y. *Rapid Communications in Mass Spectrometry* **1996**, *10*, 637-640.

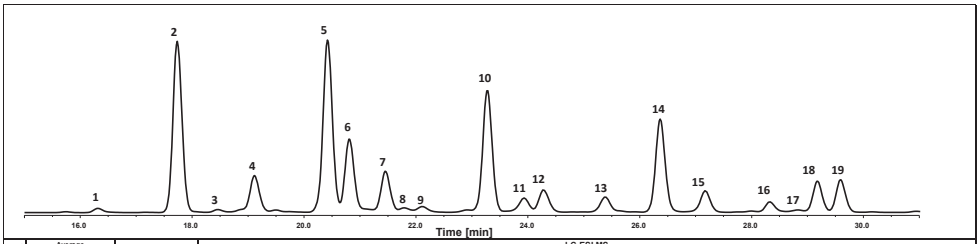
(10) Shubhakar, A.; Reiding, K. R.; Gardner, R. A.; Spencer, D. I.; Fernandes, D. L.; Wuhrer, M. *Chromatographia* **2015**, *78*, 321-333.

(11) Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. *Methods Mol. Biol* **2006**, *347*, 125-143.

(12) Klapoetke, S.; Zhang, J.; Becht, S.; Gu, X.; Ding, X. *Journal of pharmaceutical and biomedical analysis* **2010**, *53*, 315-324.

(13) Harvey, D. J.; Merry, A. H.; Royle, L.; Campbell, M. P.; Dwek, R. A.; Rudd, P. M. *Proteomics* **2009**, *9*, 3796-3801.

# Supplementary information




| HPLC Peak Id | Average GU (Procainamide) | % Area | Structure | Composition: Hex (H) | HexNAc (N) | Fuc (F) | NeuSAc (S) | LC-ESI-MS: [M/Z]$^{+}$ calculated | [M/Z]$^{2+}$ calculated | [M/Z]$^{3+}$ calculated | [M/Z]$^{+}$ registered | [M/Z]$^{2+}$ registered | [M/Z]$^{3+}$ registered | [M/Z] characteristic fragment ions (composition) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 5.29 | 0.50 | | 3 | 4 | 0 | 0 | 1536.67 | 768.84 | 512.89 | 1536.87 | 768.96 | nd | 1463.62 (H3N4-PROC)* 968.45 (H2N2-PROC)<br>1333.68 (H3N3-PROC) 895.46 (H2N2-PROC)*<br>1260.59 (H3N3-PROC)* 806.46 (H1N2-PROC)<br>1171.60 (H2N3-PROC) 644.41 (N2-PROC)<br>1130.57 (H2N2-PROC)* 441.30 (N1-PROC) |
| 2 | 5.64 | 18.62 | | 3 | 4 | 1 | 0 | 1682.73 | 841.87 | 561.58 | 1682.80 | 841.99 | nd | 1609.75 (H3N4F1-PROC)* 1317.84 (H2N3F1-PROC) 952.70 (H1N2F1-PROC) 587.53 (N1F1-PROC)<br>1536.86 (H3N4-PROC) 1276.79 (H3N2F1-PROC) 880.69 (H1N2F1-PROC)* 571.43 (N2-PROC)*<br>1480.09 (H3N3F1-PROC) 1171.77 (H2N3-PROC) 790.66 (N2F1-PROC) 514.41 (N1F1-PROC)*<br>1406.78 (H3N3F1-PROC)* 1114.76(H2N2F1-PROC) 717.51 (N2F1-PROC)* 441.42 (N1-PROC)<br>1333.80 (H3N3-PROC) 1042.75(H2N2F1-PROC)* 644.49 (N2-PROC) |
| 3 | 5.83 | 0.33 | | 5 | 2 | 0 | 0 | 1453.61 | 727.81 | 485.54 | 1454.67 | 727.92 | nd | 1014.41 (H5N1) 1014.41 (H5N1) [646.90]$^{2+}$ (H4N2-PROC) [403.84]$^{3+}$ (H1N2-PROC)<br>852.34 (H4N1) 852.34 (H4N1) 528.28 (H2N1) 366.27 (H1N1)<br>806.46 (H1N2-PROC) 806.46 (H1N2-PROC) [565.88]$^{2+}$ (H3N2-PROC) [322.81]$^{3+}$ (N2-PROC)<br>690.27 (H3N1) 690.27 (H3N1) [484.85]$^{2+}$ (H2N2-PROC) 204.11 (N1)<br>644.44 (N2-PROC) 644.44 (N2-PROC) 441.40 (N1-PROC) |
| 4 | 6.01 | 4.73 | | 3 | 5 | 1 | 0 | 1885.80 | 943.41 | 629.27 | nd | 943.52 | nd | 1682.85 (H3N4F1-PROC) 1317.71 (H2N3F1-PROC) [740.40]$^{2+}$ (H3N3F1-PROC) 441.40 (N1-PROC)<br>1536.72 (H3N4-PROC) 1299.53 (H3N4) 644.47 (N2-PROC) 366.18 (H1N1)<br>1520.74 (H2N4F1-PROC) 1098.45 (H3N3) [659.30]$^{2+}$ (H2N3F1-PROC) 350.25 (N1F1)<br>1479.74 (H3N3F1-PROC) [841.92]$^{2+}$ (H3N4F1-PROC) 587.42 (N1F1-PROC) 204.08 (N1)<br>1333.78 (H3N3-PROC) 790.41 (N2F1-PROC) 569.25 (H2N1) |
| | | | | 4 | 4 | 0 | 0 | 1698.72 | 849.86 | 566.91 | nd | 849.95 | nd | 1495.77 (H4N3-PROC) 1055.47 (H4N2) 528.31 (H2N1)<br>1333.73 (H3N3-PROC) 969.55 (H2N2-PROC) 441.33 (N1-PROC)<br>1258.60 (H4N3) [748.42]$^{2+}$ (H4N3-PROC) 366.19 (H1N1)<br>1171.84 (H2N3-PROC) [667.37]$^{2+}$ (H3N3-PROC) 204.07 (N1)<br>1130.63 (H3N2-PROC) 644.38 (N2-PROC) |
| 5 | 6.37 | 19.07 | | 4 | 4 | 1 | 0 | 1844.78 | 922.89 | 615.60 | nd | 922.97 | nd | 1641.79 (H4N3F1-PROC) 1276.61 (H3N2F1-PROC) 952.59 (H2N3) 528.22 (H2N1)<br>1495.72 (H4N3-PROC) 1276.61 (H4N3) 893.41 (H3N2) 441.32 (N1-PROC)<br>1479.73 (H3N3F1-PROC) 1132.56 (H3N2-PROC) 690.28 (H3N1) 366.13 (H1N1)<br>1334.63 (H3N3-PROC) 1114.57 (H2N2F1-PROC) 644.56 (N2-PROC) 204.08 (N1)<br>1317.64 (H2N3F1-PROC) 1114.57 (H3N3) 587.34 (N1F1-PROC) |
| | | | | 4 | 5 | 0 | 0 | 1901.80 | 951.40 | 634.60 | nd | 951.50 | nd | 1698.80 (H4N4-PROC) 1258.47 (H4N3) 644.34 (N2-PROC)<br>1536.76 (H3N4-PROC) 1171.61 (H2N3-PROC) 528.19 (H2N1)<br>1461.64 (H4N4) [849.90]$^{2+}$ (H4N4-PROC) 441.28 (N1-PROC)<br>1374. 76 (H2N4-PROC) [768.80]$^{2+}$ (H3N4-PROC) 366.13 (H1N1)<br>1338.68 (H3N3-PROC) 690.31 (H3N1) 204.06 (N1) |
| 6 | 6.48 | 8.58 | | 4 | 4 | 1 | 0 | 1844.78 | 922.90 | 615.60 | nd | 922.97 | nd | 1642.79 (H4N3F1-PROC) 1276.61 (H3N2F1-PROC) 969.53 (H2N2-PROC) 528.20 (H2N1)<br>1495.69 (H4N3-PROC) 1276.61 (H4N3) 952.56 (H2N3) 441.22 (N1-PROC)<br>1479.74 (H3N3F1-PROC) 1130.66 (H3N2-PROC) 893.38 (H3N2) 366.15 (H1N1)<br>1334.81 (H3N3-PROC) 1130.66 (H2N2F1-PROC) 644.56 (N2-PROC) 204.03 (N1)<br>1317.73 (H2N3F1-PROC) 1114.60 (H3N3) 587.34 (N1F1-PROC) |
| 7 | 6.67 | 4.73 | | 4 | 5 | 1 | 0 | 2047.86 | 1024.43 | 683.29 | nd | 1024.51 | nd | 1845.87 (H4N4F1-PROC) 1461.63 (H4N4) 441.32 (N1-PROC)<br>1682.83 (H3N4F1-PROC) 1317.70 (H2N3F1-PROC) 366.16 (H1N1)<br>1642.79 (H4N3F1-PROC) 1259.48 (H4N3)<br>1520.76 (H2N4F1-PROC) 587.40 (N1F1-PROC)<br>1479.74 (H3N3F1-PROC) 528.08 (H2N1) |
| 8 | 6.77 | 0.56 | | 6 | 2 | 0 | 0 | 1616.67 | 808.84 | 539.56 | nd | 808.93 | nd | 1292.63 (H4N2-PROC) [727.92]$^{2+}$ (H5N2-PROC) 528.28 (H2N1)<br>1176.46 (H6N1) 690.28 (H3N1) 441.37 (N1-PROC)<br>1130.63 (H3N2-PROC) [646.88]$^{2+}$ (H4N2-PROC) 366.20 (H1N1)<br>1014.40 (H5N1) 644.44 (N2-PROC) [322.76]$^{3+}$ (N2-PROC)<br>852.35 (H4N1) [565.88]$^{2+}$ (H3N2-PROC) 204.13 (N1) |
| 9 | 6.87 | 0.79 | | 5 | 4 | 0 | 0 | 1860.77 | 930.89 | 620.93 | nd | 931.00 | 621.05 | 1657.78 (H5N3-PROC) 1130.59 (H3N2-PROC) 441.33 (N1-PROC)<br>1495.82 (H4N3-PROC) 1056.22 (H4N2 366.17 (H1N1)<br>1420.61 (H5N3) 968.54 (H2N2-PROC) 204.06 (N1)<br>1333.72 (H3N3-PROC) 644.36 (N2-PROC)<br>1258.53 (H4N3) 528.26 (H2N1) |

**Supplementary Table 1.** Structural characterization of procainamide labelled IgG glycans. Human IgG glycans were released and labelled on the liquid-handling workstation as described in the Methods Section, followed by analysis by LC-ESI-MS/MS.The following symbols are used to depict glycan structures:[13] ◊, galactose; ♦, *N*-acetylgalactosamine; ○, mannose; ◈, fucose; ■, *N*-acetylglucosamine; ★, *N*-acetylneuraminic acid; dashed line, α-linkage; solid line, β-linkage.

Glycan compositions are given in the terms of hexose (H), *N*-acetylhexosamine (N), deoxyhexose (F), *N*-acetylneuraminic acid (S). * Mass corresponds to o loss of diethylamine ion (73 Da).

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **10** | **7.22** | **14.04** | | 5 | 4 | 1 | 0 | 2006.83 | 1003.40 | 669.62 | nd | 1003.99 | nd | 1641.88 (H4N3F1-PROC)<br>1495.72 (H4N3-PROC)<br>1479.75 (H3N3F1-PROC)<br>1420.59 (H5N3)<br>1333.63 (H3N3-PROC) | 1317.78 (H2N3F1-PROC)<br>1276.76 (H3N2F1-PROC)<br>1218.50 (H5N2)<br>1114.63 (H2N2F1-PROC)<br>952.83 (H1N2F1-PROC) | 790.48 (N2F1-PROC)<br>587.42 (N1F1-PROC)<br>528.18 (H2N1)<br>441.31 (N1-PROC)<br>366.16 (H1N1) | | |
| **11** | **7.43** | **1.80** | | 5 | 5 | 1 | 0 | 2209.91 | 1105.46 | 737.31 | nd | 1105.57 | 737.44 | 1845.85 (H4N4F1-PROC)<br>1698.81 (H4N4-PROC)<br>1682.81 (H3N4F1-PROC)<br>1641.75 (H4N3F1-PROC)<br>1624.65 (H5N4) | 1539.69 (H3N4-PROC)<br>1479.78 (H3N3F1-PROC)<br>1420.63 (H5N3)<br>1333.59 (H3N3-PROC)<br>1258.51 (H4N3) | 1318.75 (H2N3F1-PROC)<br>1157.47 (H1N3F1-PROC)<br>952.72 (H1N2F1-PROC)<br>790.47 (N2F1-PROC)<br>587.40 (N1F1-PROC) | 528.16 (H2N1)<br>441.30 (N1-PROC)<br>366.15 (H1N1) | |
| **12** | **7.54** | **2.66** | | 4 | 4 | 1 | 1 | 2135.87 | 1068.44 | 712.63 | nd | 1068.54 | 713.07 | 1642.74 (H4N3F1-proc)<br>1479.71 (H3N3F1-PROC)<br>1334.72 (H3N3)-PROC0<br>1317.68 (H2N3F1-PROC)<br>1277.59 (H3N2F1-PROC) | 1131.53 (H3N2-PROC)<br>1114.60 (H2N2F1-PROC)<br>[867.48]$^{[2]}$ (H4N3F1S1-PROC<br>[813.94]$^{[2]}$ (H4N4F1-PROC)<br>[886.4]$^{[2]}$ (H3N3F1S1-PROC | 819.37 (H2N1S1)<br>657.30 (H1N1S1)<br>644.38 (N2-PROC)<br>587.38 (N1F1-PROC)<br>441.33 (N1-PROC) | 366.17 (H1N1) | |
| **13** | **7.90** | **1.84** | | 5 | 4 | 0 | 1 | 2151.87 | 1076.44 | 717.96 | nd | 1076.54 | 718.09 | 1786.84 (H4N3S1-PROC)<br>1729.85 (H5N3S1)<br>1549.66 (H4N3S1)<br>1495.75 (H4N3-PROC)<br>1420.64 (H5N3) | 1333.68 (H3N3-PROC)<br>1130.56 (H3N2-PROC)<br>968.53 (H2N2-PROC)<br>819.33 (H2N1S1)<br>657.31 (H1N1S1) | 644.44 (N2-PROC)<br>528.28 (H2N1)<br>441.30 (N1-PROC)<br>366.19 (H1N1) | | |
| **14** | **8.23** | **10.51** | | 5 | 4 | 1 | 1 | 2297.93 | 1149.47 | 766.65 | nd | 1149.55 | nd | 1932.89 (H4N3F1S1-PRO<br>1787.78 (H4N3S1-PROC)<br>1770.72 (H3N3F1S1-PRO<br>1711.67 (H5N3S1)<br>1641.78 (H4N3F1-PROC) | 1508.69 (H5N2S1)<br>1479.75 (H3N3F1-PROC)<br>1333.80 (H3N3-PROC)<br>1317.56 (H2N3F1-PROC)<br>1276.68 (H3N2F1-PROC) | 1184.47 (H3N2S1)<br>1130.91 (H3N2-PROC)<br>1114.53 (H2N2F1-PROC)<br>[1077.05]$^{[2]}$ (H5N4S1-PROC<br>981.31 (H3N1S1) | [967.45]$^{[2]}$ (H4N3F1S1-PRO<br>952.34 (H1N2F1-PROC)<br>[923.44]$^{[2]}$ (H4N4F1-PROC)<br>819.28 (H2N1S1)<br>790.31 (N2F1-PROC) | 657.30 (H1N1S1)<br>587.39 (N1F1-PROC)<br>441.27 (N1-PROC)<br>366.17 (H1N1) |
| **15** | **8.52** | **2.71** | | 5 | 5 | 1 | 1 | 2501.01 | 1251.01 | 864.34 | nd | 1251.05 | nd | 2136.01 (H4N4F1S1-PRO<br>1989.88 (H4N4S1-PROC)<br>1974.89 (H3N4F1S1-PRO<br>1932.83 (H4N3F1S1-PRO<br>1844.87 (H4N4F1-PROC) | 1770.38 (H3N3F1S1-PROC)<br>1712.64 (H5N3S1)<br>1682.84 (H3N4F1-PROC)<br>1642.77 (H4N3F1-PROC)<br>1536.56 (H3N4-PROC) | 1479.73 (H3N3F1-PROC)<br>1317.75 (H2N3F1-PROC)<br>1277.69 (H3N2F1-PROC)<br>1114.91 (H2N2F1-PROC)<br>819.28 (H2N1S1) | 790.38 (N2F1-PROC)<br>657.26 (H1N1S1)<br>587.35 (N1F1-PROC)<br>441.18 (N1-PROC)<br>366.12 (H1N1) | |
| **16** | **8.94** | **1.16** | | 5 | 4 | 0 | 2 | 2442.96 | 1221.99 | 814.99 | nd | 1222.02 | nd | 2002.76 (H5N3S2)<br>1786.87 (H4N3S1-PROC)<br>1624.76 (H3N3S1-PROC)<br>1496.73 (H4N3-PROC)<br>1333.65 (H3N3-PROC) | [1077.00]$^{[2]}$ (H5N4S1-PROC)<br>819.35 (H2N1S1)<br>657.28 (H1N1S1)<br>644.44 (N2-PROC)<br>441.32 (N1-PROC) | 366.12 (H1N1) | | |
| **17** | **9.13** | **0.30** | | 5 | 5 | 0 | 2 | 2646.04 | 1323.53 | 882.69 | nd | nd | 882.78 | 1989.89 (H4N4S1-PROC)<br>1828.13 (H3N4S1-PROC)<br>1786.91 (H4N3S1-PROC)<br>1711.66 (H5N3S1)<br>1699.82 (H4N4-PROC) | 1549.61 (H4N3S1)<br>1536.56 (H3N4-PROC)<br>1495.68 (H4N3-PROC)<br>1387.55 (H3N3S1)<br>1333.68 (H3N3-PROC) | [1222.54]$^{[2]}$ (H5N4S2-P)<br>[1103.42]$^{[2]}$ (H5N4S2)<br>1171.64 (H2N3-PROC)<br>[996.96]$^{[2]}$ (H4N4S1-PROC)<br>969.59 (H2N2-PROC) | [914.94]$^{[2]}$ (H3N4S1-PROC)<br>819.34 (H2N1S1)<br>[769.39]$^{[2]}$ (H3N4-PROC)<br>657.30 (H1N1S1)<br>644.44 (N2-PROC) | 528.28 (H2N1)<br>441.35 (N1-PROC)<br>366.18 (H1N1)<br>292.11 (S1)<br>204.05 (N1) |
| **18** | **9.26** | **3.42** | | 5 | 4 | 1 | 2 | 2589.02 | 1295.01 | 863.68 | nd | 1295.06 | 864.13 | 1933.91 (H4N3F1S1-PRO<br>1787.91 (H4N3S1-PROC)<br>1770.85 (H3N3F1S1-PRO<br>1712.66 (H5N3S1)<br>1641.79 (H4N3F1-PROC) | 1508.66 (H5N3S1)<br>1479.75 (H3N3F1-PROC)<br>1276.68 (H3N2F1-PROC)<br>[1222.50]$^{[2]}$ (H5N4S2-PROC)<br>[1140.58]$^{[2]}$ (H5N4F1S1-PRO | [1088.54]$^{[2]}$ (H4N4F1S1-PRO<br>[867.94]$^{[2]}$ (H4N3F1S1-PRO<br>819.34 (H2N1S1)<br>790.44 (N2F1-PROC)<br>657.29 (H1N1S1) | 644.56 (N2-PROC)<br>587.44 (N1F1-PROC)<br>441.26 (N1-PROC)<br>366.14 (H1N1) | |
| **19** | **9.43** | **6.35** | | 5 | 5 | 1 | 2 | 2792.10 | 1396.55 | 931.37 | nd | nd | 931.43 | 1549.55 (H4N3S1)<br>1479.76 (H3N3F1-PROC)<br>1388.56 (H3N3S1)<br>1333.66 (H3N3-PROC)<br>1317.69 (H2N3F1-PROC) | [1295.57]$^{[2]}$ (H5N5S2-PROC)<br>[1251.07]$^{[2]}$ (H5N5F1S1-PRO<br>1171.33 (H2N3-PROC)<br>[1149.48]$^{[2]}$ (H5N5S1-PROC)<br>[1088.54]$^{[2]}$ (H4N4F1S1-PRO | [887.98]$^{[2]}$ (H3N4F1S1-PRO<br>819.35 (H2N1S1)<br>790.53 (N2F1-PROC)<br>657.31 (H1N1S1)<br>644.47 (N2-PROC) | 587.40 (N1F1-PROC)<br>441.42 (N1-PROC)<br>366.18 (H1N1)<br>292.15 (S1)<br>204.09 (N1) | |

**Supplementary Table 1.** (Continued).

# Chapter 6

## Variation of human salivary O-glycome

**Radoslaw P. Kozak[1], Paulina A. Urbanowicz[1], Chamindie Punyadeera[2], Karli R. Reiding[3], Bas C. Jansen[3] Louise Royle[1], Daniel I. Spencer[1], Daryl L. Fernandes[1] and Manfred Wuhrer[3]**

[1]Ludger Ltd., Culham Science Centre, Oxfordshire, OX14 3EB, United Kingdom

[2]School of Biomedical Sciences, Institute of Health and Biomedical Innovations, Queensland University of Technology, 60 Musk Avenue, Kelvin Grove, Australia

[3]Centre for Proteomics and Metabolomics Leiden University Medical Centre, The Netherlands

PLOS ONE (2016) 1-15.

## Abstract

The study of saliva O-glycosylation is receiving increasing attention due to the potential of glycans for disease biomarkers, but also due to easy access and non-invasive collection of saliva as biological fluid. Saliva is rich in glycoproteins which are secreted from the bloodstream or produced by salivary glands. Mucins which are highly O-glycosylated proteins are particularly abundant in human saliva. Their glycosylation is associated with blood group and secretor status, and represents a reservoir of potential disease biomarkers. This study aims to analyse and compare O-glycans released from whole human mouth saliva collected 3 times a day from a healthy individual over a 5 days period. O-linked glycans were released by hydrazinolysis, labelled with procainamide and analysed by ultra-high performance liquid chromatography with fluorescence detection (UHPLC-FLR) coupled to electrospray ionization mass spectrometry (ESI-MS/MS). The sample preparation method showed excellent reproducibility and can therefore be used for biomarker discovery. Our data demonstrates that the O-glycosylation in human saliva changes significantly during the day. These changes may be related to changes in the salivary concentrations of specific proteins.

## Introduction

Saliva is a unique fluid that lubricates and buffers the oral cavity.[1,2] It protects the oral cavity from microbial attack involving both the innate and acquired immune surveillances. Saliva is a structurally very complex secretion containing components from a variety of glands such as the parotid, submaxillary, sublingual and submandibular glands and some minor glands in the lip, cheek, tongue and palate.[3] Saliva contains a wide spectrum of microbes, epithelial cells, proteins/peptides, electrolytes, hormones, nasal and bronchial secretions and serum product.[4-7]

Saliva contains many glycoproteins, some of which are highly glycosylated (such as MUC5B and MUC7) as well as salivary agglutinin, secretory immunoglobulin A (sIgA), lactoferrin, amylase and proline-rich glycoproteins.[8] Mucins are dominant compounds in human saliva and play an important role in the maintenance of oral health and protection of teeth as part of the first line oral defence.[9,10] Many glycoproteins bind to a variety of bacteria and either aid or prevent adherence of bacteria to mucosal and tooth surfaces or lubricate tooth surfaces and possibly protect proteins from proteolytic attack.[11]

Bacteria usually have a net negative surface charge and therefore tend not to bind directly to negatively charged mucins. Instead, they may interact through bridging reagents involving divalent cations such as $Ca^{2+}$ ions. However, many bacteria express cell surface proteins (lectins) that specifically recognize oligosaccharides present on the salivary mucins. In this manner salivary mucins are bound by multiple bacteria resulting in agglutination. Removal of the terminal sialic acids from the oligosaccharide side chains of salivary mucins was found to abolish agglutination of oral streptococci, indicating that the bacterial lectin has specificity for sialic acids.[12] Salivary agglutinin, a highly glycosylated protein frequently associated

with other salivary proteins and with secretory immunoglobulin A (IgA), is one of the main salivary components responsible for bacteria agglutination.[13]

Secretory IgA is the most abundant humoral immunologic component of saliva. It can neutralize viruses, bacteria, and enzyme toxins. It binds bacterial antigens and is able to aggregate bacteria and inhibit their adherence to oral tissues.[14,15] Other immunologic components, such as IgG and IgM, occur in lower quantity and probably originate from gingival fluid.[16]

Salivary composition shows associations with various oral and systemic diseases [17-22], and significant age or sex related differences have been reported with respect to genomic [23,24] and proteomic expression levels.[25-27]

Recently interest in glycomics (the study of the spectrum of glycans expressed in a biological system) has increased due to the potential clinical utility of glycans as disease biomarkers.[28] Glycosylation of proteins is a very complex and common post-translational modification influencing many biological functions. There are two main types of glycosylation found in eukaryotes: N-glycosylation and O-glycosylation. On salivary glycoproteins, N-glycans are structurally diverse encompassing the high mannose, hybrid and complex N-glycan subclasses. The O-glycans comprise various O-glycan core types, with optionally fucoses in different linkages as well as sialylation and sulphation.[29,30] It has been reported that salivary MUC5B contains O-glycans which vary between individuals and that some of this variation is related to blood group and secretor status.[31] Also O-glycans attached to salivary MUC7 showed changes in the patients with recurrent aphthous stomatitis (RAS). Patients with RAS have less complex O-glycans suggesting that there are bacterial oligosaccharide degrading enzymes reducing glycan complexity.[32]

Human saliva proteomics has proven to be a novel approach in the search for

protein biomarkers for detection of diseases [31], including cancer, autoimmune diseases, viral diseases, bacterial diseases, cardiovascular diseases, and human immunodeficiency virus (HIV).[2,33] It has become clear that proteomic composition of saliva can reflect the pathological as well as physiological state.[34,35] Recently, a growing number of proof-of-principle assays have been established using saliva to monitor diseases or bodily conditions such as infection, immune responses to viral infections, systemic levels of drugs, and the detection of illicit drug use.[36] Saliva profiling over the course of disease progression could reveal potential biomarkers indicative of different stages of diseases, which may be useful in medical diagnostics. This follows from several studies showing that glycosylation patterns in other body fluids such as plasma correlate with progression of certain diseases including inflammatory conditions and subclasses of diabetes.[37-39]

Here we describe the evaluation of longitudinal changes in salivary O-glycosylation profiles using samples collected 3 times a day over a 5 days period. The O-glycans were released by hydrazinolysis using a recently developed improved protocol [40] and fluorescently labelled with procainamide, followed by HILIC-UHPLC-FLR-ESI-MS analysis. The chosen techniques proved to be suitable to demonstrate major changes in the salivary O-glycome during the day.

## Materials and methods

### Materials

All reagents and kits for O-glycan release, labelling and clean-up were from Ludger Ltd. (Oxford, UK). Centrifugal filter devices with a molecular weight cut-off (MWCO) membrane of 10 kDa were obtained from Fisher Scientific UK (Leicestershire, UK). Acetonitrile (Romil; 190 SpS for UV/ gradient quality) was obtained from Charlton

Scientific (Charlton, Oxon, UK). Fetuin glycoprotein was from Ludger Ltd (Oxford, UK) and all other reagents were obtained from Sigma-Aldrich (Dorset, UK). Procainamide labelling was performed on a Hamilton Microlab STARLet liquid-handling robotic platform.

## Collection of saliva

Whole unstimulated saliva (5 mL) was collected 3 times a day over a 5 days period from two human donors. Volunteers were asked to rinse their mouths with water for 30 seconds to remove any food particles. Donors were also asked to retain from eating 2 h prior to collection [3]. PBS supplemented with protease inhibitor cocktail (2 mL, AEBSF at 2 mM, Aprotinin at 0.3 µM, Bestatin at 116 µM, E-64 at 14 µM, Leupeptin at 1 µM and EDTA at 1 mM) and antibiotic-antimycotic mixture containing 10,000 units penicillin, 10 mg streptomycin and 25 µg amphotericin B per mL were added to each saliva sample. The saliva samples were clarified by centrifugation at 3000 x g at 4°C for 20 min to remove cellular debris. The supernatant was concentrated [29] using centrifugal filter device (10 kDa MWCO membrane). Samples were then stored at -20°C until required for further experimentation.

## Release of O-glycans by hydrazinolysis

The saliva samples and fetuin glycoprotein (used as a positive control) were buffer exchanged into 0.1% TFA prior to hydrazinolysis as described previously [40]. Samples were transferred to glass vials and dried down for 16 h by vacuum centrifugation prior to the addition of hydrazine.

Fetuin and saliva samples were incubated with 450 µL anhydrous hydrazine (Ludger) at 60°C for 6 h.[41,42] Excess hydrazine was removed by centrifugal

evaporation. The samples were placed on ice for 20 min (0°C) and were re-*N*-acetylated by addition of 0.1 M sodium bicarbonate solution (200 µL) and acetic anhydride (21 µL). Samples were cleaned up by passing them through the Ludger Clean CEX cartridges (Ludger). The O-glycans were eluted off the cartridges using water (3 x 0.5 mL). Eluates were dried by vacuum centrifugation prior to fluorescent labelling.

## Procainamide labelling

Off the released O-glycans, 40% was fluorescently labelled with procainamide as described previously [43] using Ludger procainamide glycan labelling kit containing 2-picoline borane. Briefly, samples in 10 µL of water were mixed with 10 µL of procainamide labelling solution and incubated at 65°C for 1 h. The samples were evaporated to dryness by vacuum centrifugation and re-suspended in water (100 µL) for further analysis.

## LC-ESI-MS and MS/MS analysis

Procainamide labelled samples were analysed by HILIC-UHPLC-ESI-MS with fluorescence detection. Samples were injected in 10% water/90% acetonitrile (injection volume 25 µL) onto an ACQUITY UPLC® BEH-Glycan 1.7 µm, 2.1 x 150 mm column at 60°C on a Ultimate 3000 UHPLC instrument with a fluorescence detector (λex = 310nm, λem = 370nm; Thermo Scientific). The running conditions used were: Solvent A was 50 mM ammonium formate pH 4.4 made from Ludger Stock Buffer (Ludger), and solvent B was acetonitrile. Gradient conditions were: 0 to 10 min, 90% B; 10 to 95 min, 90 to 57 % B at flow rate 0.4 mL/min; 95 to 98 min, 90 to 0 % B at a flow rate of 0.4 mL/min; 98 to 99 min, 0 % B at a flow rate of 0.25

mL/min; 99 to 100 min, 0% B at a flow rate of 0.25 mL/min. 100 to 103 min, 90% B at a flow rate 0.25 mL/min; 103 to 115 min, 90% B at a flow rate 0.4 mL/min. The UHPLC system was coupled on-line to an Amazon Speed ETD (Bruker Daltonics, Bremen, Germany) with the following settings: source temperature 250°C, gas flow 10 L/min; Capillary voltage 4500 V; ICC target 200,000; maximum accumulation time 50 ms; rolling average 2; number of precursors ions selected 3, release after 0.2 min; Positive ion mode; Scan mode: enhanced resolution; mass range scanned, 200-1600; Target mass, 900. A glucose homopolymer ladder (Ludger), labelled with procainamide was used as a system suitability standard as well as an external calibration standard for GU allocation.[44]

## Sample variation

Using in-house developed scripts, raw HILIC chromatograms for 2 individuals were exported from Chromeleon 7.1 (Dionex, USA) and aligned by applying the least variance linear fit through the highest data points observed within ± 0.25 min window around retention times 22.33, 27.13, 28.70, 29.28, 32.43 and 33.64 min. Then, for each chromatogram independently, the areas of the 7 most abundant peaks were integrated centering at 20.55 min (within a ± 0.5 min window), 22.26 (± 0.35), 23.99 (± 0.35), 24.74 (± 0.5), 27.17 (± 0.35), 32.39 (± 0.25) and 33.57 (± 0.25). The obtained peak areas were normalized to the total area sum within each chromatogram, and expressed in stripplot format using R version 3.1.2 (R Foundation for statistical computing, Vienna Austria).

### Ethics statement

The Research Ethic Committee (NHS Health Research Authority, Surrey Borders REC London, UK) granted approval for this study with all participants giving written, informed consent (REC 16/LO/0883; IRAS project ID: 203515).

## Results and discussion

### Evaluation of method repeatability

The release and recovery of O-glycans still remains very challenging because there is no universal enzyme available for O-glycan release from glycoproteins. Therefore, the current best methods for universal O-glycan release rely on rather harsh chemistry. Over the last two decades a number of chemical methods for O-glycan removal have been developed, including reductive and non-reductive β-elimination and hydrazinolysis.[45] It has been reported that the most suitable method for universal O-glycan release in their non-reduced form in high yields is hydrazinolysis [41]. We decided to use this technique to release O-glycans from human saliva because the released O-glycans contain free reducing termini allowing further derivatization with fluorescent tags and analysis by HPLC-based methods.[40,46]

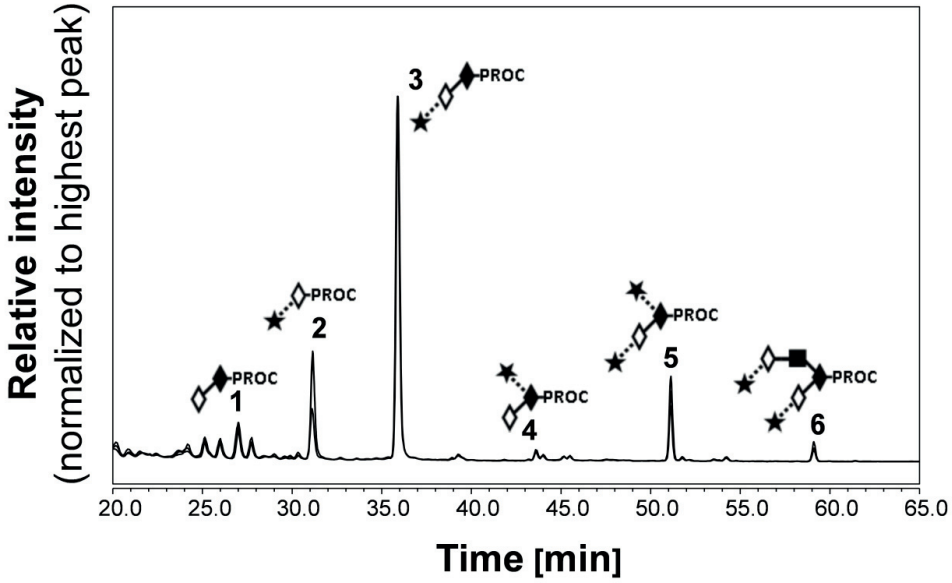



**Figure 1.** HILIC-UHPLC profiles for procainamide labelled O-glycans released in triplicate from fetuin..The following symbols are used to depict glycan structures:[47] open diamond = galactose; closed diamond = N-acetylgalactosamine; closed square = N-acetylglucosamine; closed star = N-acetyleneuraminic acid; dashed line, α-linkage; solid line, β-linkage.

First, we investigated whether the sample preparation and glycan release method is suitable for saliva samples. Using fetuin glycoprotein and a saliva reference sample collected from a single individual as positive controls, we evaluated the O-glycan preparation and O-mode hydrazinolysis. Triplicates of bovine fetuin and saliva samples were buffer exchanged into 0.1% TFA prior to release.[40] The O-glycans were released using anhydrous hydrazine, labelled with procainamide and analysed by HILIC-UHPLC with fluorescence detection providing good glycan separation, coupled to ESI-MS and MS/MS providing glycan composition.

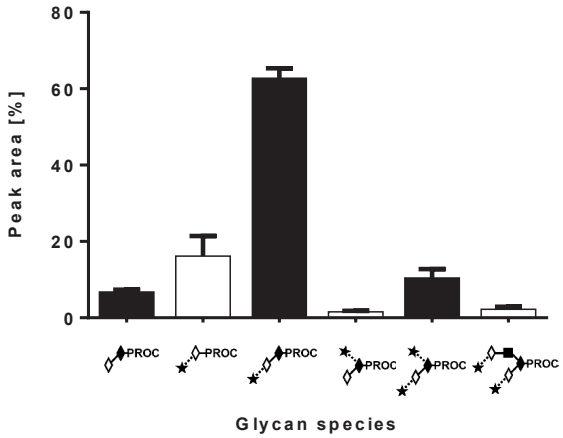



**Figure 2.** O-glycans patterns obtained from bovine fetuin. Samples were washed with 0.1% TFA followed by hydrazinolysis, procainamide labelling and HILIC-UHPLC analysis with fluorescence detection. The average abundance of the O-glycans was determined on multiple days. Symbols are as in Figure 1.

The HILIC-UHPLC profiles for triplicate releases of O-glycans from fetuin sample (Figure 1) were consistent with previously published data [46,48]. The fetuin profiles contain 6 main peaks corresponding to: core 1, Galβ1–3GalNAc (peak 1); monosialylated core 1 O-glycans, Neu5Acα2–3Galβ1–3GalNAc (peak 3); and Neu5Acα2–6(Galβ1–3)GalNAc (peak 4); disialylated core 1 O-glycan, Neu5Acα2–3Galβ1–3(Neu5Acα2–6)GalNAc (peak 5); and disialylated core 2 O-glycan, Neu5Acα2–3Galβ1–3(Neu5A- cα2–3Galβ1–4GlcNAcβ1–6)GalNAc (peak 6). The peeled product Neu5Acα2–3Gal (peak 2, Figure 1) was also detected. To confirm the identity of observed peaks (peaks 1-6, Figure 1) the procainamide labelled O-glycans were characterized by ESI-MS/MS. The results from HILIC-UHPLC coupled to ESI-MS/MS are summarised in Supplementary Table 1 and an example of fragmentation analysis of procainamide labelled O-glycans is shown in Supplementary Figure 1. The glycan peak areas for each release were very similar with the standard deviation between the samples below 5.5 % for all peaks (Figure

2).

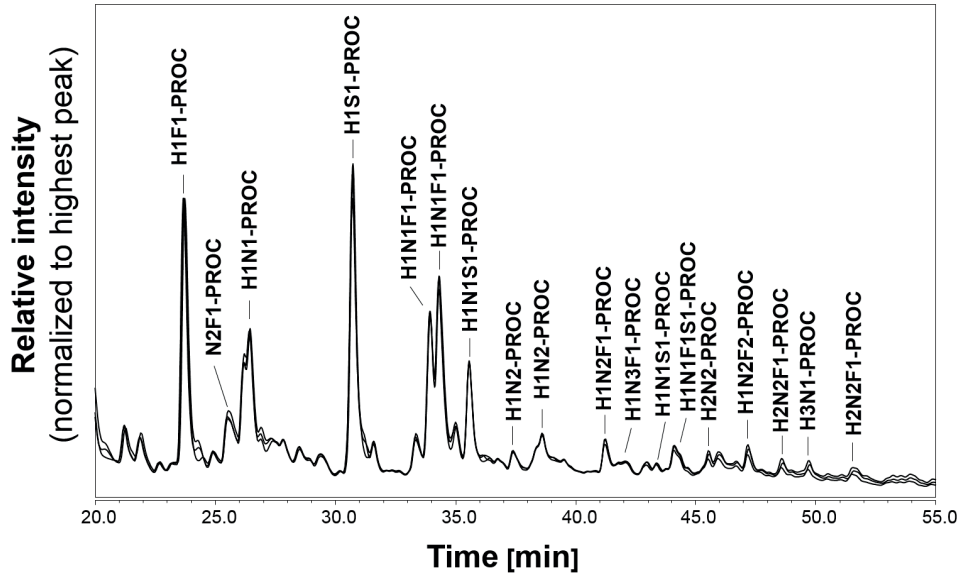



**Figure 3.** HILIC-UHPLC O-glycan profiles of saliva reference sample collected from a single individual in the morning. Sample was buffer exchange into 0.1% TFA, and the O-glycans were released in triplicate by hydrazinolysis followed by procainamide labelling and HILIC-UHPLC analysis with fluorescence detection.

The robustness of the O-glycosylation analysis method was also tested on the saliva standard. The O-glycans were released by anhydrous hydrazine, labelled with procainamide and analysed by HILIC-UHPLC-ESI-MS with fluorescence detection. The HILIC-UHPLC profiles for triplicate releases of O-glycans from whole saliva sample are shown in Figure 3. As the profiles were very complex we selected a subset of 18 peaks for comparison. These peaks were analysed by ESI-MS/MS to determine O-glycan composition. The glycan peak areas for selected peaks were very similar with the standard deviation between the samples below 1 % (Figure 4). Thus, O-mode hydrazinolysis, procainamide labelling systems and UHPLC-ESI-MS/MS analysis are highly reproducible and suitable for the analysis of human saliva samples.

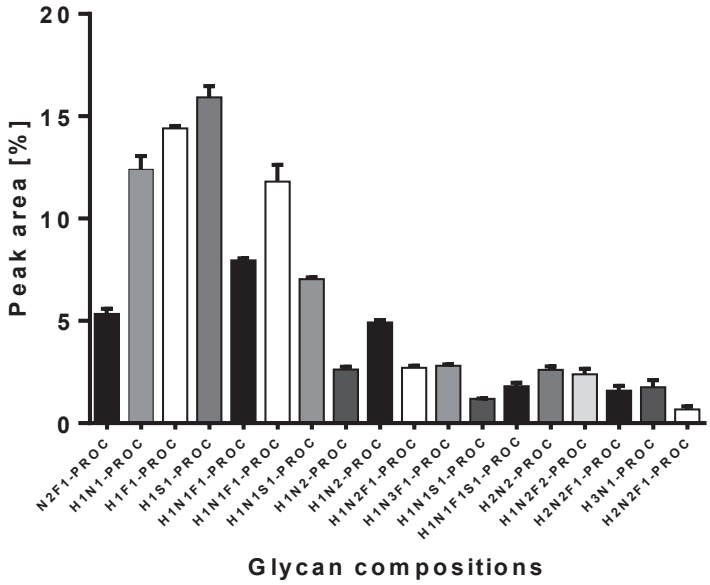



**Figure 4.** O-glycan patterns obtained from triplicate releases from saliva sample collected in the morning. Samples were washed with 0.1% TFA followed by hydrazinolysis, procainamide labelling and HILIC-UHPLC analysis with fluorescence detection. 18 O-glycan structures were selected for comparison.

## Longitudinal analysis of O-glycans from human saliva

O-glycans were released from saliva collected 3 times a day over a 5 days period from 2 individuals by hydrazinolysis after acidic wash, labelled with procainamide and analysed by UHPLC-ESI-MS/MS. Only one sample was taken through ESI-MS/MS analysis. As the obtained HILIC profiles were very complex we selected a subset of 27 peaks for comparison and detailed analysis by ESI-MS/MS (Supplementary Figure 2). A broad range of O-glycan compositions were detected in the samples and although there was some overlap in LC traces, the sample profiles showed prominent differences. The observed molecular masses and compositions of 27 selected structures (Figure 5) are summarised in Table 1.

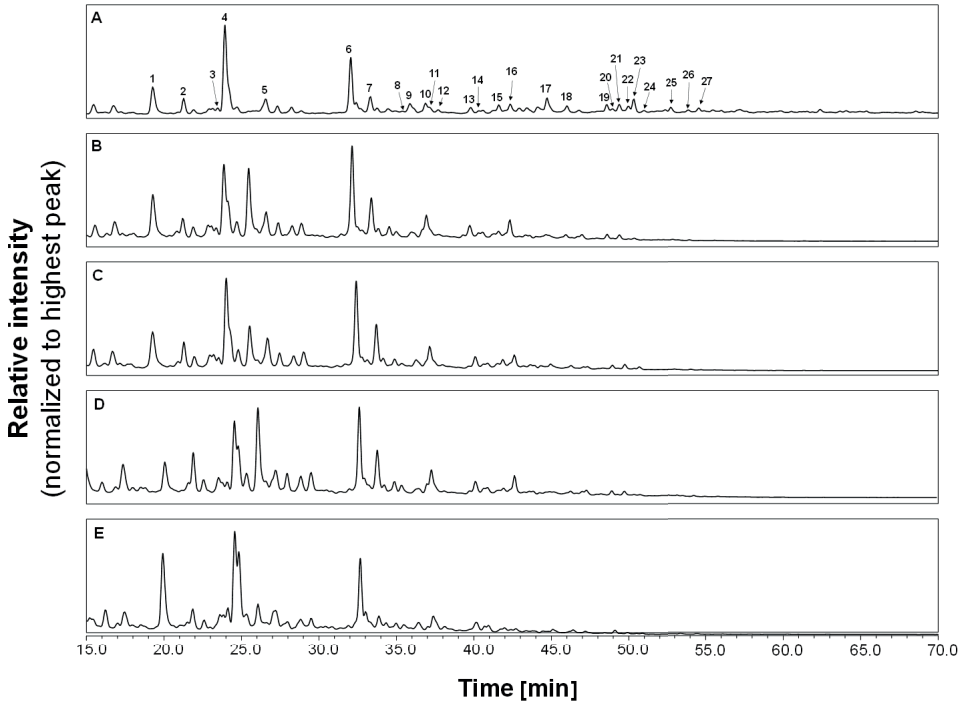



**Figure 5.** HILIC-UHPLC profile of saliva O-glycan patterns. Samples were collected in days 1-5 in the morning. Samples were buffer exchanged into 0.1% TFA, and the O-glycans were released by hydrazinolysis followed by procainamide labelling and HILIC-UHPLC analysis with fluorescence detection.

The O-glycans displayed a heterogenous mixture of neutral and acidic oligosaccharides and most of the structures were fucosylated. The MS/MS fragmentation data of procainamide labelled O-glycans contained diagnostic fragment ions characteristic of core type or blood group epitope consistent with previously reported data.[29,31,49,50] The O-glycans carried the blood group H-epitopes Fucα1-2Galβ1 (diagnostic structure: Fuc-Hex) which is usually modified by a single monosaccharide residue (Galα1-3- or GalNAcα1-3-) and determines ABO blood groups.[31,51,52] The O-glycans found in this study were released from whole saliva from an individual of blood group A. The terminal structures of glycans 13 and 18 (Figure 5, Table 1) have a HexNAc attached to (Fuc)-Hex-HexNAc-Proc (blood

group H-epitope), suggesting expression of blood group A antigen GalNAcα1-3(Fucα1-2)Galβ1- on saliva glycoproteins. The presence of ABO blood group specific epitopes matching the blood group status suggest that the saliva donor is secretor.

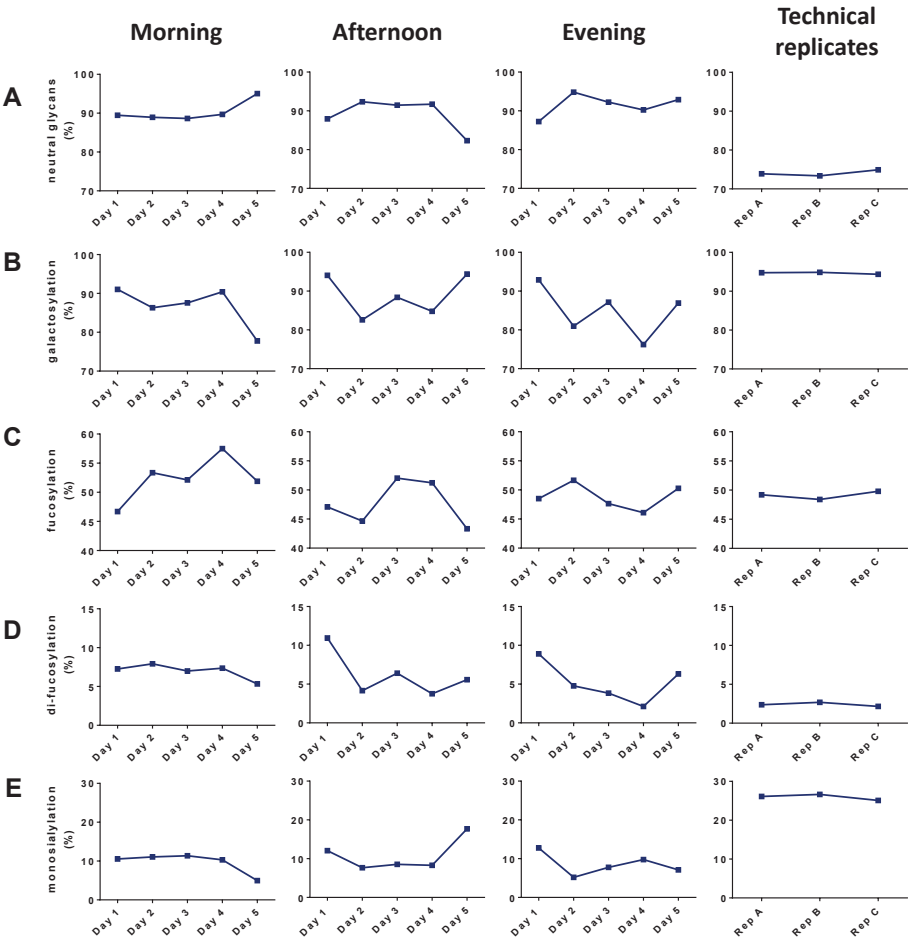



**Figure 6.** Saliva O-glycosylation variation during collection. Several differences in saliva O-glycosylation between different collection times and days can be observed with respect to (A) neutral glycans, (B) galactosylation, (C) fucosylation, (D) difucosylation, (E) monosialylation. Data for the technical replicates corresponds to the saliva reference sample collected from a single individual and analysed during evaluation of method repeatability and robustness.

| UHPLC Peak Id | Average GU (PROC) | HILIC-UHPLC-ESI-MS | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Composition | | | | Calculated $[m/z]^+$ | Registered $[m/z]^+$ | Probable structures |
| | | Hex (H) | HexNAc (N) | Fucose (F) | Neu5Ac (S) | | | |
| 1 | 1.40 | 0 | 1 | 1 | 0 | 587.38 | 587.32 | - |
| 2 | 1.51 | 1 | 0 | 1 | 0 | 546.38 | 546.30 | F-H-PROC |
| 3 | 1.63 | 0 | 2 | 1 | 0 | 791.45 | 790.41 | F-N-N-PROC |
| 4 | 1.66 | 1 | 1 | 0 | 0 | 603.39 | 603.32 | H-N-PROC |
| 5 | 1.81 | 1 | 1 | 1 | 0 | 749.45 | 749.38 | F-H-N-PROC |
| 6 | 2.21 | 1 | 1 | 1 | 0 | 749.45 | 749.38 | F-H-N-PROC |
| 7 | 2.30 | 1 | 1 | 0 | 1 | 894.48 | 894.42 | S-H-N-PROC |
| 8 | 2.45 | 1 | 2 | 0 | 0 | 806.46 | 806.40 | N-(H)-N-PROC |
| 9 | 2.48 | 1 | 2 | 0 | 0 | 806.46 | 806.40 | N-(H)-N-PROC |
| 10 | 2.57 | 1 | 1 | 2 | 0 | 895.5 | 895.44 | S-(S-H)-N-ROC |
| 11 | 2.61 | 1 | 2 | 1 | 0 | 952.51 | 952.46 | N-(S-H)-N-ROC |
| 12 | 2.64 | 1 | 3 | 0 | 0 | 1010.53 | 1009.48 | H-N-(N)-N-ROC |
| 13 | 2.82 | 1 | 2 | 1 | 0 | 952.51 | 952.46 | N-(F)-H-N-ROC |
| 14 | 2.87 | 2 | 2 | 0 | 0 | 969.5 | 968.46 | - |
| 15 | 2.98 | 1 | 1 | 0 | 1 | 894.48 | 894.42 | S-(H)-N-PROC |
| 16 | 3.05 | 1 | 3 | 1 | 0 | 1156.59 | 1155.54 | S-H-N-(N)-N-PROC |
| 17 | 3.28 | 2 | 2 | 0 | 0 | 968.52 | 968.46 | H-(H-N)-N-PROC |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 18 | 3.41 | 1 | 2 | 2 | 0 | 1098.59 | 1098.52 | N-(F)-H-(F)-N-PROC |
| 19 | 3.69 | 3 | 1 | 0 | 0 | 927.1 | 927.43 | - |
| 20 | 3.73 | 2 | 2 | 1 | 0 | 1114.59 | 1114.51 | H-N-(F-H)-N-PROC |
| 21 | 3.78 | 2 | 2 | 1 | 0 | 1114.59 | 1114.51 | H-N-(F-H)-N-PROC |
| 22 | 3.85 | 2 | 2 | 0 | 1 | 1259.64 | 1259.55 | S-H-N-H-N PROC |
| 23 | 3.89 | 2 | 2 | 1 | 0 | 1114.59 | 1114.51 | H-N-(F-H)-N-PROC |
| 24 | 3.97 | 2 | 2 | 1 | 0 | 1114.58 | 1114.51 | N-H-N-(F)-H-PROC |
| 25 | 4.18 | 3 | 2 | 0 | 0 | 1130.58 | 1130.51 | - |
| 26 | 4.31 | 2 | 2 | 2 | 0 | 1260.65 | 1260.57 | F-H-N-(F-H)-N-PROC |
| 27 | 4.40 | 2 | 2 | 1 | 1 | $[703.00]^{2+}$ | $[703.31]^{2+}$ | F-H-N-(S-H)-N-PROC |

**Table 1.** Mass spectrometric data for selected 27 O-glycans obtained from human saliva.

Derived glycosylation features such as galactosylation and fucosylation were calculated for the selected O-glycan peaks to analyse the range of variability. Firstly we compared levels of galactosylation (Figure 6B). A pronounced decrease in galactosylation (from 91% on the first day to 78% on the fifth day) was observed for the samples collected in the morning. For the samples collected in the afternoon or evening variation was also high (between 83% and 94% for afternoon collection and 76% and 93% for evening collection).

Next we looked at changes in fucosylation for the samples collected three times a day during a 5 days period (Figure 6C). High biological variation was observed both intra-day and inter-day with fucosylation levels ranging from 43% to 57%.

We also looked at changes in sialylation. Significant variation (between 5% and 13%) between samples collected at different points of the day during the 5 days period was observed (Figure 6E).

The peak alignment across all HILIC chromatograms from two individuals shows that the individuals differ in relative peak abundance, while displaying similar variation within each peak (Supplementary Figure 3C.).

The main aim of this study was to analyse and compare entire O-glycan populations released form whole human saliva samples collected 3 times a day over a 5 days period. Our results highlighted that there is extensive natural variation in the salivary O-glycosylation. Both the intra-day and the inter-day variation appeared to be high.

We used whole human saliva because the wide spectrum of compounds (including glycoproteins) present in saliva may provide information for potentially important biomarkers and it has been reported that 30% of blood proteins are also present in saliva.[2] Also changes in O-glycans have been associated with different disease states.[53,54] It has been previously published that the variation in salivary MUC5B glycosylation is high and dependent on blood group or secretor status.[31] Likewise, one may expect that salivary O-glycosylation, along with dynamic changes in composition of other salivary analytes may depend on diet, lifestyle and inflammatory responses in the mucus layer, albeit additional studies will be necessary to study this. One of the biggest advantages of saliva is easy collection, storage and transportation. Donation is non-invasive, stress-free and it does not require highly-trained personnel.[6]

## Conclusions

Our study showed that the composition and relative abundance of the O-glycans in human saliva significantly change during the day. This is an interesting observation, that should be taken into consideration when evaluating salivary O-glycosylation e.g. for its diagnostic biomarker potential.

The changes in the specific O-glycan types within one day as well as between days may be caused by changes in the salivary concentrations of specific proteins, but may likewise be attributed to changes in protein-specific glycosylation profiles. It may, therefore make sense to follow the current study up and study the glycosylation of one of the mucins abundantly present in saliva with respect to its longitudinal robustness and biomarker potential. Mucins found in saliva have high turnover rates and respond quickly to altered conditions within the body.[32,53] Mucins and their O-glycans affect a range of aspects of oral health and their changes in the response to the infection and inflammation.[32]

## Acknowledgements

This work was supported by the European Union's Seventh Framework Program (FP7-Health-F5-2011) under grant agreement n°278535 (HighGlycan).

Ludger Ltd is a commercial bioscience company specialising in development and validation of glycoprofiling technology. The following authors declare competing financial interest(s): RPK, PAU, LR, DIS, DLF are employed by Ludger Ltd. In addition, products from Ludger were used in this research. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

## Reference List:

(1) Punyadeera, C.; Dimeski, G.; Kostner, K.; Beyerlein, P.; Cooper-White, J. *J. Immunol. Methods* **2011**, *373*, 19-25.

(2) Schulz, B. L.; Cooper-White, J.; Punyadeera, C. K. *Crit Rev. Biotechnol* **2013**, *33*, 246-259.

(3) Mohamed, R.; Campbell, J. L.; Cooper-White, J.; Dimeski, G.; Punyadeera, C. *Clin. Transl. Med* **2012**, *1*, 19.

(4) Hofman, L. F. *J. Nutr* **2001**, *131*, 1621S-1625S.

(5) Kaufman, E.; Lamster, I. B. *Crit Rev. Oral Biol. Med* **2002**, *13*, 197-212.

(6) Pfaffe, T.; Cooper-White, J.; Beyerlein, P.; Kostner, K.; Punyadeera, C. *Clin. Chem* **2011**, *57*, 675-687.

(7) Sondej, M.; Denny, P. A.; Xie, Y.; Ramachandran, P.; Si, Y.; Takashima, J.; Shi, W.; Wong, D. T.; Loo, J. A.; Denny, P. C. *Clin. Proteomics* **2009**, *5*, 52-68.

(8) FML, A.; RMP, V.; MJC, I.; PMDN, D., 2007, pp 347-376.

(9) Frohm, N. M.; Sandstedt, B.; Sorensen, O.; Weber, G.; Borregaard, N.; Stahle-Backdahl, M. *Infect. Immun* **1999**, *67*, 2561-2566.

(10) Helmerhorst, E. J.; Oppenheim, F. G. *J. Dent. Res* **2007**, *86*, 680-693.

(11) Amerongen, A. V.; Bolscher, J. G.; Veerman, E. C. *Glycobiology* **1995**, *5*, 733-740.

(12) R.J., L.; H.J., J. *Oral Microbiology at a Glance*; Wiley-Blackwell, 2010.

(13) Amerongen, A. V.; Veerman, E. C. *Oral Dis* **2002**, *8*, 12-22.

(14) Schenkels, L. C.; Veerman, E. C.; Nieuw Amerongen, A. V. *Crit Rev. Oral Biol. Med* **1995**, *6*, 161-175.

(15) AR, T. C. *Oral Histology: Development, Structure And Function*, 1998.

(16) Edgar, W. M. *Br. Dent. J* **1992**, *172*, 305-312.

(17) Borzym-Kluczyk, M.; Radziejewska, I.; Cechowska-Pasko, M. *Clin. Chem. Lab Med* **2011**, *49*, 885-889.

(18) Dowling, P.; Wormald, R.; Meleady, P.; Henry, M.; Curran, A.; Clynes, M. *J. Proteomics* **2008**, *71*, 168-175.

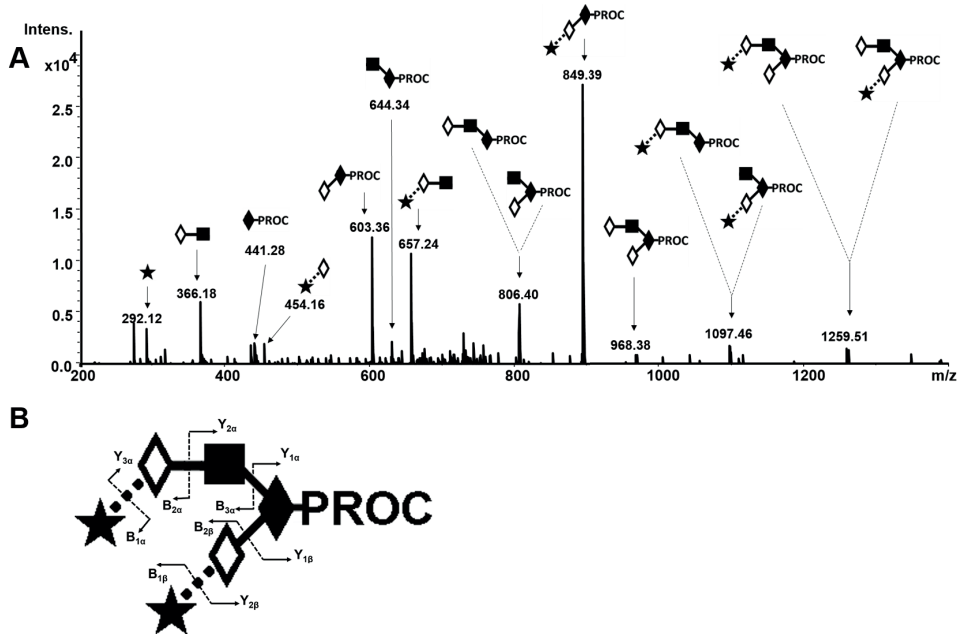
(19) Foo, J. Y.; Wan, Y.; Kostner, K.; Arivalagan, A.; Atherton, J.; Cooper-White, J.; Dimeski, G.; Punyadeera, C. *PLoS. One* **2012**, *7*, e48452.

(20) Hu, S.; Arellano, M.; Boontheung, P.; Wang, J.; Zhou, H.; Jiang, J.; Elashoff, D.; Wei, R.; Loo, J. A.; Wong, D. T. *Clin. Cancer Res* **2008**, *14*, 6246-6252.

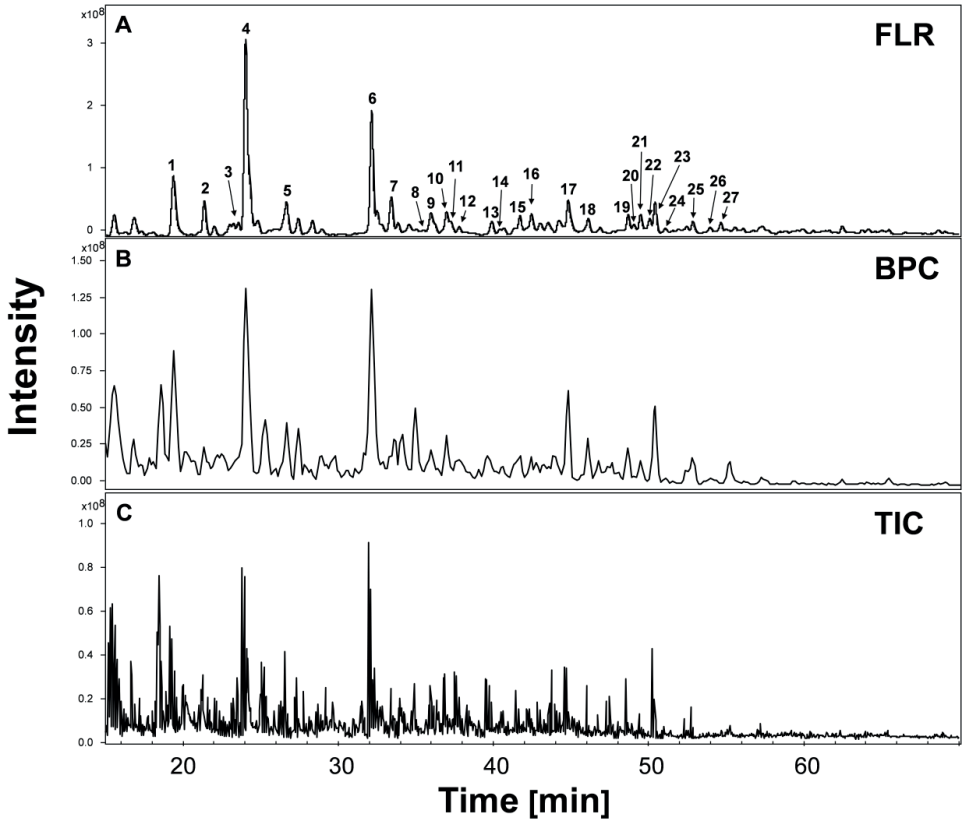
(21) Ovchinnikov, D. A.; Wan, Y.; Coman, W. B.; Pandit, P.; Cooper-White, J. J.; Herman, J. G.; Punyadeera, C. *Biomark. Insights* **2014**, *9*, 53-60.
(22) Zhang, L.; Xiao, H.; Karlan, S.; Zhou, H.; Gross, J.; Elashoff, D.; Akin, D.; Yan, X.; Chia, D.; Karlan, B.; Wong, D. T. *PLoS. One* **2010**, *5*, e15573.
(23) Bocklandt, S.; Lin, W.; Sehl, M. E.; Sanchez, F. J.; Sinsheimer, J. S.; Horvath, S.; Vilain, E. *PLoS. One* **2011**, *6*, e14821.
(24) Srivastava, A.; Wang, J.; Zhou, H.; Melvin, J. E.; Wong, D. T. *Arch. Oral Biol* **2008**, *53*, 1058-1070.
(25) Ambatipudi, K. S.; Lu, B.; Hagen, F. K.; Melvin, J. E.; Yates, J. R. *J. Proteome. Res* **2009**, *8*, 5093-5102.
(26) Cabras, T.; Pisano, E.; Boi, R.; Olianas, A.; Manconi, B.; Inzitari, R.; Fanali, C.; Giardina, B.; Castagnola, M.; Messana, I. *J. Proteome. Res* **2009**, *8*, 4126-4134.
(27) Fleissig, Y.; Reichenberg, E.; Redlich, M.; Zaks, B.; Deutsch, O.; Aframian, D. J.; Palmon, A. *Oral Dis* **2010**, *16*, 831-838.
(28) Gornik, O.; Pavic, T.; Lauc, G. *Biochim. Biophys. Acta* **2012**, *1820*, 1318-1326.
(29) Everest-Dass, A. V.; Jin, D.; Thaysen-Andersen, M.; Nevalainen, H.; Kolarich, D.; Packer, N. H. *Glycobiology* **2012**, *22*, 1465-1479.
(30) Everest-Dass, A. V.; Abrahams, J. L.; Kolarich, D.; Packer, N. H.; Campbell, M. P. *J. Am. Soc. Mass Spectrom* **2013**, *24*, 895-906.
(31) Thomsson, K. A.; Schulz, B. L.; Packer, N. H.; Karlsson, N. G. *Glycobiology* **2005**, *15*, 791-804.
(32) Zad, M.; Flowers, S. A.; Bankvall, M.; Jontell, M.; Karlsson, N. G. *Clin. Oral Investig* **2015**, *19*, 2147-2152.
(33) Hu, S.; Jiang, J.; Wong, D. T. *Methods Mol. Biol* **2010**, *666*, 31-41.
(34) Epstein, J. B.; Gorsky, M.; Guglietta, A.; Le, N.; Sonis, S. T. *Cancer* **2000**, *89*, 2258-2265.
(35) Zauber, H.; Mosler, S.; von, H. A.; Schulze, W. X. *Proteomics* **2012**, *12*, 2221-2235.
(36) Zhang, A.; Sun, H.; Wang, X. *Appl. Biochem. Biotechnol* **2012**, *168*, 1718-1727.
(37) Clerc, F.; Reiding, K. R.; Jansen, B. C.; Kammeijer, G. S.; Bondt, A.; Wuhrer, M. *Glycoconj. J* **2015**.
(38) Adamczyk, B.; Tharmalingam, T.; Rudd, P. M. *Biochim. Biophys. Acta* **2012**, *1820*, 1347-1353.

(39) Jaeken, J.; Matthijs, G. *Annu. Rev. Genomics Hum. Genet* **2007**, *8*, 261-278.
(40) Kozak, R. P.; Royle, L.; Gardner, R. A.; Fernandes, D. L.; Wuhrer, M. *Anal. Biochem* **2012**, *423*, 119-128.
(41) Merry, A. H.; Neville, D. C.; Royle, L.; Matthews, B.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M. *Anal. Biochem* **2002**, *304*, 91-99.
(42) Patel, T.; Bruce, J.; Merry, A.; Bigge, C.; Wormald, M.; Jaques, A.; Parekh, R. *Biochemistry* **1993**, *32*, 679-693.
(43) Kozak, R. P.; Tortosa, C. B.; Fernandes, D. L.; Spencer, D. I. *Anal. Biochem* **2015**, *486*, 38-40.
(44) Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. *Methods Mol. Biol* **2006**, *347*, 125-143.
(45) Zauner, G.; Kozak, R. P.; Gardner, R. A.; Fernandes, D. L.; Deelder, A. M.; Wuhrer, M. *Biol. Chem* **2012**, *393*, 687-708.
(46) Kozak, R. P.; Royle, L.; Gardner, R. A.; Bondt, A.; Fernandes, D. L.; Wuhrer, M. *Anal. Biochem* **2014**, *453*, 29-37.
(47) Harvey, D. J.; Merry, A. H.; Royle, L.; Campbell, M. P.; Dwek, R. A.; Rudd, P. M. *Proteomics* **2009**, *9*, 3796-3801.
(48) Spiro, R. G.; Bhoyroo, V. D. *J. Biol. Chem* **1974**, *249*, 5704-5717.
(49) Chaudhury, N. M.; Proctor, G. B.; Karlsson, N. G.; Carpenter, G. H.; Flowers, S. A. *Mol. Cell Proteomics* **2015**.
(50) Karlsson, N. G.; Thomsson, K. A. *Glycobiology* **2009**, *19*, 288-300.
(51) Karlsson, N. G.; Schulz, B. L.; Packer, N. H. *J. Am. Soc. Mass Spectrom* **2004**, *15*, 659-672.
(52) Schulz, B. L.; Packer, N. H.; Karlsson, N. G. *Anal. Chem* **2002**, *74*, 6088-6097.
(53) Kawashima, H. *Biol. Pharm. Bull* **2012**, *35*, 1637-1641.
(54) Schulz, B. L.; Sloane, A. J.; Robinson, L. J.; Prasad, S. S.; Lindner, R. A.; Robinson, M.; Bye, P. T.; Nielson, D. W.; Harry, J. L.; Packer, N. H.; Karlsson, N. G. *Glycobiology* **2007**, *17*, 698-712.
(55) Domon, B.; Costello, C. E. *Glycoconjugate journal* **1988**, *5*, 397-409.

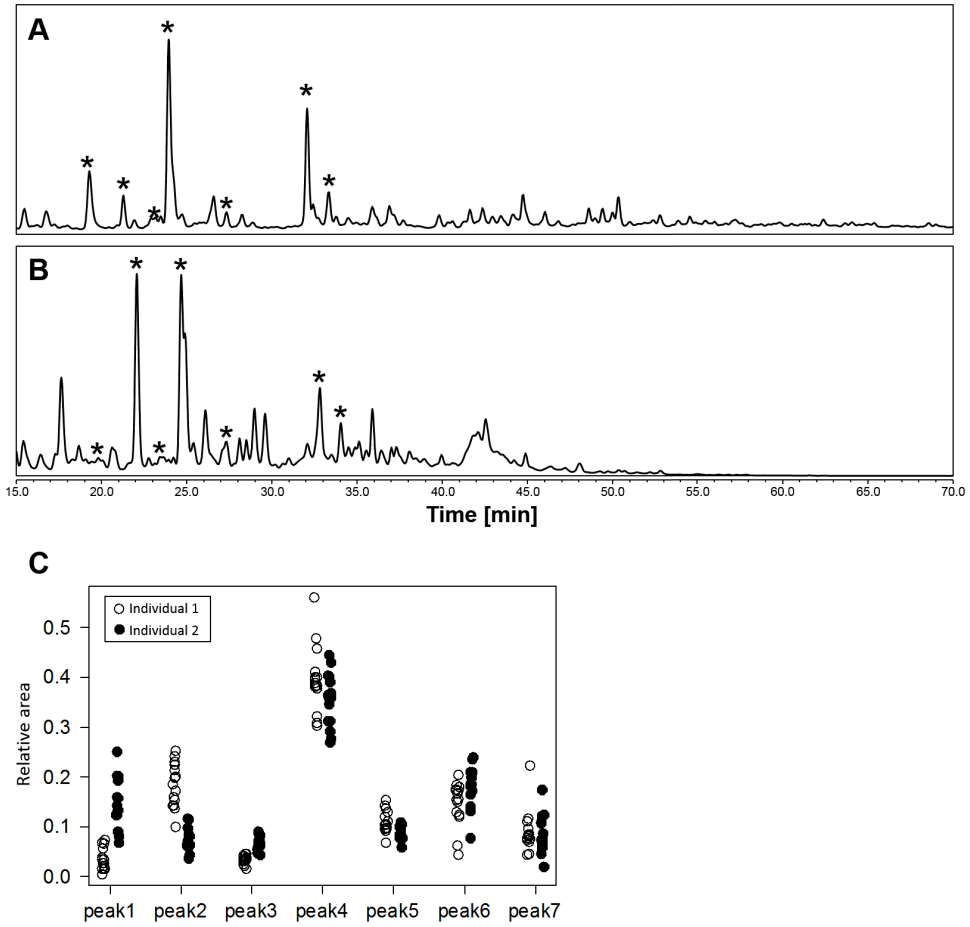
# Supplementary information




**Supplementary Figure 1.** MS/MS characterization of procainamide labelled fetuin O-glycan following online HILIC-UHPLC. (A) MS/MS characterization of the glycan eluting at retention time equivalent to that of GU 4.91. (B) Fragmentation scheme of the $[M+2H]^{2+}$ ion at m/z 775.84. The nomenclature used for fragmentation is based on literature.[55]




**Supplementary Figure 2.** Fluorescence, MS and MS/MS traces of saliva O-glycan patterns. Sample was collected in days 1 in the morning. (A) HILIC-FLR trace; (B) Base Peak Chromatogram (BPC); (C) Total Ion Chromatogram (TIC).




**Supplementary Figure 3.** Comparison of HILIC-UHPLC O-glycan profiles of saliva samples collected from two individuals. The 7 peaks labelled with asterisks correspond to signals suitable for alignment and integration. (A) Saliva sample from individual 1 collected at day 1 in the morning; (B) Saliva sample from individual 2 collected at day 1 in the morning; (C) Comparison between individual 1 and 2 of the variation across time within 7 O-glycans measured by HILIC-UHPLC.
After chromatogram alignment, the signals were integrated and the obtained areas normalized to the total area sum. The individuals show to differ in relative peak abundance, while displaying similar variation within each peak.

| UHPLC Peak Id | Average | | Structure | HILIC-UHPLC-ESI-MS and MS/MS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | GU (PROC) | % Area | | Composition | | | $[m/z]^{+}$ calculated | $[m/z]^{2+}$ calculated | $[m/z]^{+}$ registered | $[m/z]^{2+}$ registered | [m/z] characteristic B and Y fragment ions (composition) |
| | | | | Hex (H) | HexNAc (N) | Neu5Ac (S) | | | | | |
| 1 | 1.93 | 6.80 | | 1 | 1 | 0 | 603.32 | 302.17 | 603.35 | nd | 441.30 (N1-PROC) |
| 2 | 2.18 | 16.15 | | 1 | 0 | 1 | 691.34 | 346.17 | 691.36 | nd | 292.12 (S); 400.27 (H1-PROC) |
| 3 | 2.50 | 62.82 | | 1 | 1 | 1 | 894.42 | 447.71 | 894.41 | nd | 441.31 (N1-PROC); 603.34 (H1N1-PROC) |
| 4 | 3.13 | 1.59 | | 1 | 1 | 1 | 894.42 | 447.71 | 894.41 | nd | 441.32 (N1-PROC); 603.37 (H1N1-PROC) |
| 5 | 3.90 | 10.42 | | 1 | 1 | 2 | 1185.51 | 593.26 | 1185.54 | 593.30 | 292.13 (S); 441.26 (N1-PROC); 454.16 (HS); 603.33 (H1N1-PROC); 732.36 (N1S1-PROC); 894.38 (H1N1S1-PROC) |
| 6 | 4.91 | 2.22 | | 2 | 2 | 2 | 1550.65 | 775.83 | nd | 775.84 | 292.12 (S); 366.18 (H1N1); 441.28 (N1-PROC); 454.16 (H1S1); 603.36 (H1N1-PROC); 644.34 (N2-PROC); 657.24 (H1N1S1); 806.40 (H1N2-PROC); 894.39 (H1N1S1-PROC); 968.38 (H2N2-PROC); 1097.46 (H1N2S1-PROC); 1259.51 (H2N2S1-PROC) |

**Supplementary Table 1.** Summary of O-glycan structures from fetuin with GU values, average relative abundance and mass values.

# Chapter 7

## General discussion

In the 1960s, glycans were viewed as unimportant structures. Carbohydrates attached to the proteins were seen as components that might increase protein solubility or could contribute to the protection against proteolysis.[1] Glycans were not considered for further functional significance. We now know this not to be true at all. From today's perspective, it might be concluded that those views were caused by the lack of analytical techniques, platforms, and also accurate data.

Over the last decades the attention given by researchers, biopharmaceutical companies and regulators to glycosylation has greatly increased.[2,3] However, the progress made was very unequal in that most of the technology has been developed for the characterisation of N-glycans, leaving analysis of the O-glycans behind. In the following, the status of glycoanalytical techniques and our knowledge of glycan function will be compared for protein N- and O-glycosylation.

Over the last years several methods have been developed for the characterisation of N-glycans. The main methods involve removal of N-glycans from glycoproteins by enzymatic digestion with PNGase F, a labelling with fluorescent tag and subsequent separation using chromatography methods. The structures then can be confirmed by exoglycosidase sequencing and analysis by HPLC, mass spectrometry or other orthogonal methods. Significant progress has been made with regards to the development of high-throughput HPLC and MS technologies for the analysis of N-glycans in large sample sets,[4] This glycomics strategy allows the identification of glycosylation changes during different stages of the disease,

including monitoring of progression, and the discovery of glycan-based biomarkers.[5]

N-glycans have been also shown to influence the safety and efficacy of biological drugs. The Glycosylation Critical Quality Attributes (GCQAs) such as sialylation, core fucosylation, antennary composition, N-acetyllactosamine extensions, alpha-galactose and N-glycolylneuraminic acid need to be optimised and controlled during drug development.[6]

In contrast, much less is understood about O-glycans which makes our overall knowledge of protein glycosylation still rather incomplete. The O-glycans are attached to the hydroxyl groups in serine and threonine residues and include O-linked GlcNAc, mannose, fucose, and GalNAc (mucin type). The O-glycosylation is one of the most complex regulated posttranslational modifications of the proteins and also one of the least understood.

The work described in this thesis is focused on mucin type O-glycosylation. This type of O-glycosylation is present on >80 % of proteins that pass through the secretory apparatus and has been found to be important in many normal and pathological settings.[7,8]

Unlike for N-glycans, the biosynthesis of O-glycans occurs by sequential addition of single monosaccharides rather than via the initial transfer of a common conserved lipid-linked precursor to the growing protein backbone which allows for the enormous variability of O-glycan attachments to proteins in different cells (it is difficult to pinpoint their exact locations). O-glycans can also display multiple core structures and several of these O-linked cores might be present on a single protein.[9,10] Analysis of mucin type O-glycosylation has

been challenging because of this diversity and complexity, with the result that our understanding of O-glycosylation lags behind.

Recent studies indicate that alteration of mucin type O-glycosylation leads to diverse effects and changes in O-glycosylation have been associated with different diseases, such as IgA nephropathy, tumour formation and metastasis, cancer or Tn syndrome.[11-14] The O-linked GalNAc (Tn antigen), the core 1 O-glycan Galα1-3GalNAc (T antigen) and the mono-sialylated core 1 O-glycan Neu5Acα2-3Galβ1-3GalNAc (STn antigen) have been found to be abundantly expressed in adenocarcinomas and thus are considered as tumour-associated antigens. Mucins containing these truncated O-glycans were found to be immunogenic in humans especially when conjugated to MUC1 peptide backbones [15] , thus raising the possibility of using the "cancer-like" MUC1 glycopeptides as cancer vaccines. However, despite these studies our knowledge of the biological functions of O-glycans in the context of cancer and biologics is still very limited. Several very important questions can be asked:

1. How do O-glycan structures regulate biology?
2. How can O-glycans modulate key events of cancer cells such as growth, apoptosis, angiogenesis, and adherence?
3. What are the biotechnological implications of O-glycans?

It appears that the main reason why so little is known about O-glycans is because there is a lack of simple, sensitive and effective technologies for O-glycosylation analysis.[16]

Since O-glycans may affect structural and functional properties of carrier proteins or peptides, the eventual control of O-glycosylation would

allow one to manipulate proteins and peptides and alter their biological effects. Clearly, in view of the above mentioned interest in vaccine, biopharmaceuticals as well as in the clinical development, improved analytical methods to study O-glycans are urgently needed.

In general, O-glycosylation analysis can be performed at three different levels:

a) analysis of glycans on intact glycoproteins
b) analysis of glycopeptides and
c) structural analysis and characterisation of released glycans.

While there are numerous techniques for protein O-glycosylation analysis (as described in Chapter 2), no single one can provide all the information desired, and no method is without its limitations. Successful analysis involves the complete characterisation of the O-linked glycoprotein, including the composition and structure of the O-glycans and the structural heterogeneity at each occupied amino acid site.

As was previously mentioned, each of these developed methods has specific advantages and disadvantages. Release of O-glycans remains rather challenging (Chapter 2). The main reason for this is that, in contrast of the situation to the analysis of N-glycans there is no enzyme available for universal O-glycan release from proteins. O-glycans require chemical strategies for their release, which can be inefficient, and very often result in O-glycan degradation via a peeling reaction. Peeling is a major problem when preforming the O-glycan release and often results in poor repeatability with variable amounts of the small peeled glycans. Many biopharmaceutical developers have investigated this peeling problem. In Chapters 3 and 4 the

development of sample preparation techniques for O-glycan release (in their nonreduced form) by hydrazinolysis are described. These have been optimized to reduce the undesirable peeling reaction. A drawback of previous methods was a high level (of up to 70%) of this undesirable side reaction. In the methods described in this thesis, the peeling product was reduced to 17-20%. At the moment, the most common way for separation of released glycans is liquid chromatography (LC) and with the major separation principles being hydrophilic interaction liquid chromatography (HILIC) [17,18], weak anion exchange chromatography (WAX) [18,19], reversed-phase chromatography (RP) [18,20] or porous graphitized carbon chromatography (PGC).[21] Most of these separation approaches are suitable for coupling to mass spectrometry, making them attractive tools for O-glycan analysis.

The improved protocols developed in this thesis allow the release of O-glycans with free reducing termini and only limited degradation. The O-glycans released with the "free end" can be fluorescently labelled and analysed by several techniques including microarray technology,[22,23] Fluorescently labelled glycans can be covalently immobilised on commercial epoxide-activated glass slides via the secondary amine group that links the glycan moiety with the fluorescent tag. This technology allows multiparameter testing of a large number of samples at minimal sample levels and it is increasingly applied for the study of protein-carbohydrate interactions for viral agglutinins, human lectins, and serum anti-carbohydrate antibodies with many target glycans from biological systems. To date, glycan microarray studies have been primarily performed with synthetic glycans, or N-glycans and GSL-glycans that have been enzymatically derived from natural sources. These

arrays can also impact drug discovery by allowing the specificities of lectins, growth factors, cytokines, antibodies and toxins to be analysed rapidly.[24] The O-glycan arrays are under-represented mainly due to problems with O-glycan release. Our improved hydrazinolysis procedure combined with reductive amination can be essential for future O-glycan array studies which may this provide important information on protein- carbohydrate interactions.

Alternatively, fluorescently labelled O-glycans have been be analysed by capillary gel electrophoresis with laser-induced fluorescence (CGE-LIF).[25,26] The CGE-LIF analysis is one of the most powerful separation tools for glycan analysis providing great separation of both positional and linkage isomers. This analysis also requires glycan derivatization with a fluorescent tag, which contains the charged group needed for migration in an electric field to increase the sensitivity of the laser-induced fluorescence (LIF) detection. The most common charged fluorescent tag used for CE analysis is 8-aminopyrene-1,3,6-trisulfonate (APTS).[27,28] APTS labelled glycans can be analysed with high speed and resolution. It has been reported that major glycans found in mAbs can be analysed in less than 10 min. Also online coupling of CE with MS could help with O-glycan characterization. However, sialylated species run quickly and cannot be resolved.

In view of this, we have investigated the use of highly sensitive fluorescent tags for HILIC-HPLC and MS analysis. In Chapter 5 we have shown that procainamide labelled glycans gave a stronger fluorescent response and significantly higher ion intensity in positive ESI-MS when compared to the commonly used 2-AB. The procainamide-labelled glycans are suitable for both (U)HPLC–FLR analysis, providing good chromatographic

separation for relative quantitation, and for ESI–MS analysis, providing more efficient ionization for glycan identification. We expect that procainamide may automatically substitute 2-AB in most analytical workflows involving reductive amination of glycans. The combination of improved hydrazinolysis and procainamide labelling might be a strategy that drug developers and researchers could use for O-glycan analysis of low amounts of sample.

However, the current chemical O-glycan release methods require large amounts of biological or biopharmaceutical samples. In this context a primary need is simple and sensitive technology. Future miniaturization would be desired to fulfil the increasing demands for analysing smaller amounts of glycoprotein and higher throughput of samples. One possible way is to achieve this is the use nanoscale format columns for separation and subsequent detection by MS. The nano LC can serve as a trapping and concentration device for O-glycan species, significantly improving the sensitivity of detection. Combination of porous graphitized carbon (PGC) separation with MS is one possible way for the analysis of native or permethylated O-glycan alditols by PGC-(nano)LC-MS.[21,29] The PGC columns are highly selective and separation of the reduced O-glycans eliminates α/β-isomerization of the reducing end hydroxyl group which significantly simplify the analysis. PGC also shows great capability for the separation of structural and linkage isomers of complex glycans with varying degree of sialylation and sulphation. However this technology is not widely applied in biopharmaceutical development. One of the reasons for this is the lack of public databases of retention properties of glycans. Also the analysis of the previously mentioned Tn and STn antigens is very difficult and a recent

publication has suggested the lack of reproducibility and robustness of PGC glycan analysis.[30] This issue still has to be resolved before the PGC glycans analysis can be applied routinely.

Another promising approach might be the development of miniaturized hydrazinolysis systems for O-glycan release with free reducing termini and fluorescent labelling. Reducing the amount of anhydrous hydrazine and the reagents for re-N-acetylation and labelling could significantly reduce the sample amounts needed for analysis, time and costs. Minimizing sample handling, transfer steps, and reducing cation exchange column diameters, solvent flow-rates could greatly enhance the analytical sensitivity.

Using the strategy described in Chapters 3, 4 and 5, combining O-glycan release by improved hydrazinolysis, with procainamide labelling followed by HILIC-UHPLC-ESI-MS/MS analysis we have studied O-glycosylation of human saliva. Human saliva is a potentially rich source of clinical biomarkers and can be collected noninvasively. Saliva contains many glycoproteins, some of which are highly glycosylated (such as MUC5B and MUC7) and changes in the O-glycosylation have been associated with different disease states. It has been previously reported that the variation in salivary MUC5B glycosylation is high and dependent on blood group or secretor status. However, not much is known regarding the influence the physiological parameters such as diet, lifestyle and inflammatory responses that influence on the glycosylation patterns. Our study revealed that the composition and relative abundance of the O-glycans in human saliva significantly change during the day. The changes in the specific O-glycan types within one day as well as between days may be caused by changes in

the salivary concentrations of specific proteins, but may likewise be attributed to changes in protein-specific glycosylation profiles. It may, therefore make sense to perform follow up study which will focus on one of the mucins abundantly present in saliva with respect to its longitudinal robustness and biomarker potential.

Glycans play critical roles in in many aspects of human health. However, as mentioned before relatively little attention has been paid to O-glycans, and O-glycomics area still remains a relatively understudied field. However, it is hard to predict which advances will be made in understanding the functions of O-glycans in human health, disease and biopharmaceutical development.

An expansion of our understanding of O-glycan structures and functions and applying this knowledge could help with early detection of cancers, other diseases, vaccines and biopharmaceutical development.

In conclusion, further studies are needed to better understand and evaluate the biological functions of O-glycan and for assessing therapeutic efficacy of the mucin type O-glycosylation. Use of the improved O-glycan release techniques and procainamide labelling system described in this thesis allowed detailed analysis of O-glycans from both glycoprotein standards and complex human saliva by U(H)PLC-FLR-ESI-MS/MS. I believe that this technology could be used by researchers and drug developers for detection and quantification of biopharmaceutical O-glycans and might be a powerful tool for O-linked glycoprotein biomarker discovery.

## Reference List:


(1) Eylar, E. H. *Journal of Theoretical Biology* **1966**, *10*, 89-113.

(2) Walsh, G. *Nat. Biotechnol* **2010**, *28*, 917-924.

(3) *Biopharmaceuticals - A Global Market Overview*, 2013.

(4) Shubhakar, A.; Reiding, K. R.; Gardner, R. A.; Spencer, D. I.; Fernandes, D. L.; Wuhrer, M. *Chromatographia* **2015**, *78*, 321-333.

(5) Marino, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. *Nat. Chem. Biol* **2010**, *6*, 713-723.

(6) Fernandes, D. L. In *Biopharmaceutical Production Technology, Volume 1 & Volume 2 (ed G. Subramanian)*; Wiley-VCH Verlag GmbH & Co.KGaA,Weinheim,Germany., 2012.

(7) Schjoldager, K. T.; Clausen, H. *Biochim. Biophys. Acta* **2012**, *1820*, 2079-2094.

(8) Kudelka, M. R.; Antonopoulos, A.; Wang, Y.; Duong, D. M.; Song, X.; Seyfried, N. T.; Dell, A.; Haslam, S. M.; Cummings, R. D.; Ju, T. *Nat. Methods* **2016**, *13*, 81-86.

(9) Hounsell, E. F.; Davies, M. J.; Renouf, D. V. *Glycoconj. J* **1996**, *13*, 19-26.

(10) Yamashita, Y.; Chung, Y. S.; Horie, R.; Kannagi, R.; Sowa, M. *J. Natl. Cancer Inst* **1995**, *87*, 441-446.

(11) Allen, A. C.; Bailey, E. M.; Brenchley, P. E.; Buck, K. S.; Barratt, J.; Feehally, J. *Kidney Int* **2001**, *60*, 969-973.

(12) Ju, T.; Cummings, R. D. *Proc. Natl. Acad. Sci. U. S. A* **2002**, *99*, 16613-16618.

(13) Hiki, Y.; Tanaka, A.; Kokubo, T.; Iwase, H.; Nishikido, J.; Hotta, K.; Kobayashi, Y. *J. Am. Soc. Nephrol* **1998**, *9*, 577-582.

(14) Mall, A. S. *J. Clin. Pathol* **2008**, *61*, 1018-1024.

(15) Corfield, A. P. *Biochim. Biophys. Acta* **2015**, *1850*, 236-252.

(16) Zauner, G.; Kozak, R. P.; Gardner, R. A.; Fernandes, D. L.; Deelder, A. M.; Wuhrer, M. *Biol. Chem* **2012**, *393*, 687-708.

(17) Zauner, G.; Deelder, A. M.; Wuhrer, M. *Electrophoresis* **2011**, *32*, 3456-3466.

(18) Royle, L.; Mattu, T. S.; Hart, E.; Langridge, J. I.; Merry, A. H.; Murphy, N.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M. *Anal. Biochem* **2002**, *304*, 70-90.

(19) Guile, G. R.; Harvey, D. J.; O'Donnell, N.; Powell, A. K.; Hunter, A. P.; Zamze, S.; Fernandes, D. L.; Dwek, R. A.; Wing, D. R. *Eur. J. Biochem* **1998**, *258*, 623-656.
(20) Kuraya, N.; H
ase, S. *Anal. Biochem* **1996**, *233*, 205-211.
(21) Stavenhagen, K.; Kolarich, D.; Wuhrer, M. *Chromatographia* **2015**, *78*, 307-320.
(22) Lonardi, E.; Balog, C. I.; Deelder, A. M.; Wuhrer, M. *Expert. Rev. Proteomics* **2010**, *7*, 761-774.
(23) de Boer, A. R.; Hokke, C. H.; Deelder, A. M.; Wuhrer, M. *Anal. Chem* **2007**, *79*, 8107-8113.
(24) Liang, P. H.; Wu, C. Y.; Greenberg, W. A.; Wong, C. H. *Curr. Opin. Chem. Biol* **2008**, *12*, 86-92.
(25) Reusch, D.; Haberger, M.; Kailich, T.; Heidenreich, A. K.; Kampe, M.; Bulau, P.; Wuhrer, M. *MAbs* **2014**, *6*, 185-196.
(26) Mechref, Y. *Electrophoresis* **2011**, *32*, 3467-3481.
(27) Guttman, A. *Nature* **1996**, *380*, 461-462.
(28) Szabo, Z.; Guttman, A.; Rejtar, T.; Karger, B. L. *Electrophoresis* **2010**, *31*, 1389-1395.
(29) Jensen, P. H.; Karlsson, N. G.; Kolarich, D.; Packer, N. H. *Nat. Protoc* **2012**, *7*, 1299-1310.
(30) Melmer, M.; Stangler, T.; Premstaller, A.; Lindner, W. *J. Chromatogr. A* **2011**, *1218*, 118-123.

**English summary**

**Dutch summary**

**Abbreviations**

**Curriculum Vitae**

**List of publications**

**Acknowledgements**

## English summary

Glycans are very diverse structures and provide a valuable source of biomarkers. The O-glycans on glycoproteins and mucins play very important roles in immunity, leukocyte trafficking, vascular biology, angiogenesis and lymphangiogenesis. Mucin O-glycans also protect gut epithelial cells from the microflora and extreme pH and prevent tumorigenesis.

Changes in both N- and O-glycosylation have been associated with many states of health and disease, providing prognostic and diagnostic information. O-glycans have been relatively little studied through lack of suitable analytical tools but the evidence is that they could be promising biomarkers.

O-glycan profiles successfully distinguished breast cancer samples from normal samples. Truncated O-glycans on the hinge region of IgA are associated with IgA nephropathy. Future studies need to further define key O-glycans whose alterations reflect disease progression. Such knowledge will be great importance critical in developing novel diagnostic and therapeutic strategies.

Over the last decade the number of approved protein biopharmaceuticals has been rapidly growing with a total sales value of $140 billion representing approximately 20% of the total pharmaceutical market. Protein biopharmaceuticals include fusion proteins, cytokines, monoclonal antibodies (mAb), vaccines, hormones, growth factors, and blood factors. More than two-thirds of these recombinant biopharmaceuticals products are glycoproteins, and glycans have been associated with safety and efficacy of biopharmaceuticals. Most of the glycoprotein therapeutics bear N-glycans but O-glycans are also found in some economically important biopharmaceuticals such as EPO. The FDA technical guidance for

characterizing and monitoring glycans says: “Glycan structures should be characterized, and particular attention should be paid to their degree of mannosylation, galactosylation, fucosylation and sialylation. The distribution of main glycan structures present should be determined.”

The aim of the work described in this thesis was to develop and improve methods for O-glycosylation analysis. Over the last decade a number of techniques for O-glycan analysis have been developed, however release of O-glycans, the first essential step of analysis, remains very challenging (Chapter 2). The main reason for this is that there is no enzyme available for universal O-glycan release from proteins. The commercially available O-glycanases, or endo-α-N-acetylgalactosamidases, have a high specificity limited to the release of neutral core 1 O-glycans (Galβ1-3GalNAc). Therefore, chemical release is the most effective method for releasing the full range of O-glycans. Several chemical methods for liberation of O-glycans have been described and the most common one is reductive β-elimination. This release methods leads to O-glycans in their reduced forms (alditols). This makes it impossible to introduce reducing-end tags for fluorescent detection and the analysis is restricted to mass spectrometry, high-pH anion-exchange chromatography with pulsed amperometric detection or nuclear magnetic resonance. Several techniques for chemical release of O-glycans in non-reduced form have been developed and it has been reported that the best method for universal O-glycan release in high yields is hydrazinolysis.

Unfortunately all of the O-glycan release methods show a side reaction known as “peeling”, which is a stepwise degradation of the polysaccharide starting at the reducing end and removing one sugar residue at a time. Peeling is a well-known

problem when performing O-glycan release and often results in poor repeatability with variable amounts of the small peeled glycans. This is a major problem for comparability studies or quality control. Biopharmaceutical developers and researchers have investigated this peeling problem. In Chapters 3 and 4 the development of sample preparation techniques for O-glycan release (in their nonreduced form) by hydrazinolysis are described. The methods have been optimized to reduce the undesirable peeling reaction. Developed techniques show that buffer exchange into 0.1% TFA or 100 mM EDTA prior to hydrazinolysis or addition of EDTA directly to hydrazine significantly reduces the amount of unwanted peeling. A drawback of previous methods was a high level (up to 70%) of this side reaction. With the methods described in this thesis, peeling product went down to 17-20% making the O-glyan analysis more robust.

One of the most widely used methods for glycan analysis is reducing-end labelling of released glycans by reductive amination followed by HILIC-UHPLC with fluorescent detection. The stoichiometric attachment of one tag per glycan allows the relative quantitation of different glycan species based on fluorescence or ultraviolet (UV) absorbance intensity. Several fluorescent tags have been used for glycan derivatization such as 2-aminobenzamide (2-AB), 2-aminobenzoic acid (2-AA), 1-aminopyrene-3,6,8-trisulfonic acid (APTS), 2-aminopyridine (AP), procaine, and procainamide. In Chapter 5 we have shown that procainamide-labelled glycans gave higher fluorescent response and significantly higher ion intensity in positive ESI-MS when compared to the most widely used label 2-AB. The procainamide-labelled glycans are suitable for both (U)HPLC–FLR analysis, providing good chromatographic separation for relative quantitation, and ESI–MS analysis, providing more efficient ionization for glycan identification.

Using the strategy described in Chapters 3 (“Suppression of peeling during the release of O-glycans by hydrazinolysis”), 4 (“Improved nonreductive O-glycan release by hydrazinolysis with EDTA addition”) and 5 (“Comparison of procainamide and 2-aminobenzamide labelling for profiling and identification of glycans by LC-FLR-ESI-MS”), combining O-glycan release by hydrazinolysis with procainamide labelling followed by HILIC-UHPLC-ESI-MS/MS analysis we studied O-glycosylation of human saliva. We used human saliva because the wide spectrum of compounds (including glycoproteins) present in saliva may provide information for potentially important biomarkers and it is relatively easy to obtain. Our study showed that the composition and relative abundance of the O-glycans in human saliva significantly changed during the day. This observation should be taken into consideration by researches evaluating salivary O-glycosylation for its diagnostic biomarker potential.

The closing chapter of this thesis (Chapter 7) contains a general discussion, which places the individual chapters into context. We believe that our technology for O-glycosylation analysis can be applied by researchers for analysis of complex biological samples and for O-glycan biomarker discovery. It is also anticipated that presented, improved technology will be used by drug developers at the early stages of drug development through to post-approval batch release. This strategy can be incorporated into drug developers’ programmes to support QbD-based realisation and regulatory submissions of glycoprotein therapeutics. However, further studies are needed to better understand and evaluate biological functions or for assessing therapeutic the efficacy of mucin type O-glycosylation.

## Dutch summary

Eiwitglycosylering is een veel voorkomende post-translationele modificatie die resulteert in de vorming van zogenaamde glycoproteïnen. Bij de twee meest voorkomende typen, O- en N-glycosylering, worden zowel relatief eenvoudige als complexe glycaanstructuren aangetroffen. Het onderzoek dat in dit proefschrift wordt beschreven richt zich op O-glycosylering. O-glycanen spelen o.a. een belangrijke rol in het afweersysteem, migratie van witte bloedcellen in het lichaam en de aanmaak en functie van lymfe- en bloedvaten. Daarnaast beschermen O-glycanen van mucines op darmepitheelcellen tegen microflora en extreme pH en hebben ze een beschermende werking bij het ontstaan van tumoren.

In studies waarbij de eiwitglycosylering van zowel zieke als gezonde mensen werd onderzocht is gebleken dat veranderingen in eiwitglycosylering gecorreleerd zijn met verschillende (stadia van) ziekten. Hiermee hebben glycosyleringprofielen dus wellicht prognostische en diagnostische waarde en kunnen bepaalde glycaanstructuren zelfs mogelijk als biomarker worden gebruikt. Een voorbeeld van een succesvolle classificatie op basis van O-glycaan profielen is beschreven voor borstkanker. Daarnaast zijn verkorte O-glycanen in een specifieke structuur van het immunoglobuline A (IgA), de zogenaamde hinge region, in verband gebracht met een bepaalde vorm van nierontsteking (IgA nefropathie). Meer onderzoek is echter noodzakelijk teneinde vast te stellen welke O-glycaan structuren correleren met progressie van een ziekte. Deze kennis is essentieel voor de verdere ontwikkeling van diagnostische en therapeutische strategieën.

De afgelopen 10 jaar is het aantal goedgekeurde eiwitgeneesmiddelen, zogenaamde biologicals, snel gegroeid, wat heeft geresulteerd in een totale marktwaarde van

$140 miljard. Dat is ongeveer 20% van de totale markt van alle geneesmiddelen. Onder biologicals vallen (recombinante) eiwitten zoals cytokines, monoclonale antilichamen, vaccins, hormonen groeifactoren en bloedfactoren. Meer dan tweederde van deze biologicals zijn glycoproteïnen, en de glycanen op deze eiwitten zijn belangrijk voor de werkzaamheid, maar ook veiligheid, van deze geneesmiddelen. De FDA richtlijnen schrijven hierover: “Glycaanstructuren dienen gekarakteriseerd te worden, en speciale aandacht dient te worden gegeven aan de mate van mannosylering, galactosylering, fucosylering en sialylering. De verdeling van de belangrijkste glycaanstructuren dient te worden bepaald.”

Voor zowel de studies naar O-glycosyleringsprofielen bij ziekteprocessen als de karaktersering van O-glycanen op biologicals is de ontwikkeling van methoden en technieken die de (grootschalige) analyse van O-glycanen mogelijk maakt onontbeerlijk. Het doel van het in dit proefschrift beschreven onderzoek is de ontwikkeling en verbetering van zulke methoden.

De laatste 10 jaar zijn er verschillende technieken ontwikkeld voor de analyse van O-glycanen. Hiervoor is het als eerste belangrijk om de O-glycanen los in handen te krijgen. Echter, het verbreken van de verbinding tussen de O-glycanen en de eiwitten waaraan ze gekoppeld zitten, blijft een lastige klus (Hoofdstuk 2). De belangrijkste reden hiervoor is dat er geen universeel enzym beschikbaar is dat alle O-glycanen van eiwitten kan vrijmaken. De commercieel verkrijgbare O-glycanases, endo-α-N-acetylgalactosamidases, hebben een hoge specificiteit die zich beperkt tot zogenaamde neutrale core-1 O-glycanen (Galβ1-3GalNAc). Om die reden wordt er vaak voor een chemische reactie voor het vrijmaken van de O-glycanen gekozen. Hiervoor zijn verschillende methoden beschreven en de bekendste hiervan is β-

eliminatie. Deze strategie resulteert in O-glycanen in de gereduceerde vorm (alditol). Het nadeel van alditolen is dat deze niet aan het reducerende uiteinde gekoppeld kunnen worden met een fluorescent label waardoor de analyse van alditolen beperkt blijft tot massaspectrometrie (MS), anionenwisselingschromatografie met PAD (pulsed amperometric detection) of nucleaire resonantie spectroscopie (NMR). Om O-glycanen in een niet-reducerende vorm te verkrijgen zijn andere methoden beschikbaar. De meest geschikte voor het verkrijgen van een zo breed mogelijk spectrum van verschillende structuren met een hoge opbrengst is hydrazinolyse.

Helaas gaan de chemische methoden voor het losmaken van O-glycanen gepaard met nevenreacties, dit wordt peeling genoemd. Peeling is de stapsgewijze degradatie van een oligosacharide en begint bij het reducerende uiteinde. Het is een bekende nevenreactie die in veel gevallen heeft geresulteerd in een hoge mate van variatie in de O-glycaanprofielen die worden gemeten. Vanzelfsprekend is dit een groot probleem voor vergelijkende studies en kwaliteitscontrole en daarom is peeling door veel groepen bestudeerd. In de hoofdstukken 3 en 4 wordt de optimalisatie van methodes voor loskoppeling van O-glycanen met behulp van hydrazinolyse beschreven, met als doel de peeling zo laag mogelijk te houden. Wij vonden dat het toevoegen van 0.1% TFA of 100 mM EDTA, voorafgaand aan de hydrazinolyse, of toevoegen van EDTA aan hydrazine, resulteerde in een sterke reductie van de mate van peeling. Terwijl in de eerder gepubliceerde methodes peeling kon oplopen tot 70%, konden wij met onze strategie de hoeveelheid peeling terugbrengen tot 17-20%.

Eén van de meest gebruikte methodes voor glycaananalyse bestaat uit het labelen van het reducerende uiteinde van losgekoppelde glycanen met een fluorofoor via reductieve aminering. Structuren worden daarna gemonitord met behulp van HILIC-UHPLC en fluorescente detectie. Stoichiometrische koppeling van 1 label met elk glycaan maakt kwantificering van de verschillende structuren mogelijk. Fluorescente labels die worden gebruikt zijn 2-aminobenzamide (2-AB), 2-aminobenzoic acid (2-AA), 1-aminopyrene-3,6,8-trisulfonic acid (APTS), 2-aminopyridine (AP), procaine, en procainamide. In Hoofdstuk 5 laten wij zien dat procainamide-gelabelde glycanen een betere fluorescente respons geven en een hogere intensiteit in positieve ESI-MS, in vergelijking tot het meer algemeen gebruikte 2-AB label. De procainamide-gelabelde glycanen zijn zowel geschikt voor (U)HPLC–FLR analyse, met goede chromatografische scheiding en dus kwantificering, en voor ESI-MS analyse, met hogere ionisatie efficiëntie en dus betere identificatie.

In Hoofdstuk 6 wordt O-glycosylering van humaan speeksel bestudeerd, gebruikmakend van de geoptimaliseerde methodes zoals beschreven in de hoofdstukken 3, 4 en 5. Humaan speeksel bevat een grote verscheidenheid aan verbindingen, inclusief glycoproteïnen, en vormt daarom een interessante bron van potentiële biomarkers. Wij vonden dat zowel de compositie als de hoeveelheid O-glycosylering in humane speekselmonsters varieerde gedurende de dag. Dit is belangrijke informatie voor elk vervolgonderzoek waar O-glycosylering van speeksel wordt bestudeerd, vooral als het daarbij om het identificeren van biomarkers gaat.

In het laatste hoofdstuk van dit proefschrift wordt een algemene discussie gepresenteerd, waarbij de individuele hoofdstukken over de ontwikkeling en

optimalisering van verschillende technieken voor het bestuderen van eiwit O-glycosylering, in perspectief worden plaatst.

Het is onze overtuiging dat de hier gepresenteerde technologie kan worden ingezet voor de analyse van complexe biologische monsters en voor O-glycaan studies met als doel biomarkers te vinden. Wij verwachten ook dat deze verbeterde strategieën zullen worden toegepast in de farmaceutische industrie voor karakterisering van geneesmiddelen in zowel de vroege fase van ontwikkeling als aan het einde van de produktielijn en na goedkeuring van de regelgevers, waarbij de kwaliteit continu gemonitord dient te blijven. De hier beschreven strategieën voor O-glycaan structuuranalyses kunnen dus een belangrijke bijdrage leveren bij de ontwikkeling van een biological, en de volledige aanvraag voor goedkeuring daarvan. Er is echter wel meer onderzoek vereist voor het ophelderen en begrijpen van biologische functies van O-glycosylering in het algemeen en de therapeutische werkzaamheid van bijvoorbeeld mucine-achtige O-glycosylering in het bijzonder.

## Abbreviations

| | |
|---|---|
| 2-AA | 2-aminobenzoic acid |
| 2-AB | 2-aminobenzamide |
| AI-ECD | activated ion-electron capture dissociation |
| AP | 2-aminopyridine |
| APP | amyloid precursor protein |
| APTS | 1-aminopyrene-3,6,8-trisulfonic acid |
| BAP | biotin pentylamine |
| BSM | bovine submaxillary glands mucin |
| BSSL | bile salt-stimulated lipase |
| CDG | congenital disorder of glycosylation |
| CID | collision induced fragmentation |
| DTT | dithiothreitol |
| ECD | electron capture dissociation |
| EDTA | ethylenediaminetetraacetic acid |
| EMEA | European Medicines Evaluation Agency |
| ESI | electrospray ionization |
| ETD | electron transfer dissociation |
| Fab | fragment antigen binding |

| | |
|---|---|
| Fc | fragment cristallizable |
| FDA | US Food and Drug Administration |
| FLR | fluorescence |
| FTICR | fourier transform ion cyclotron resonance |
| Fuc | fucose |
| Gal | galactose |
| GalNAc | N-acetylgalactosamine |
| GalNAc | N-acetylgalactosamine |
| GC | gas chromatography |
| GCQAs | Glycosylation Critical Quality Attributes |
| Glc | glucose |
| GlcNAc | N-acetylglucosamine |
| GlcNAc | N-acetylglucosamine |
| Hex | hexose |
| HexNAc | N-acetylhexosamine |
| HILIC | hydrophilic interaction liquid chromatography |
| HPAEC-PAD | high pH anion exchange chromatography with pulsed amperometric detection |
| HPLC | high-performance liquid chromatography |
| ICH | International Conference on Harmonization |

| | |
|---|---|
| IgA | immunoglobulin A |
| IgG | immunoglobulin G |
| IgM | immunoglobulin M |
| IRMPD | infrared multiphoton dissociation |
| LC | liquid chromatography |
| LIF | laser-induced fluorescence |
| m/z | mass over charge ratio |
| mAb | monoclonal antibody |
| MALDI | matrix assisted laser desorption |
| Man | mannose |
| MS | mass spectrometry |
| NeuAc | N-acetylneuraminic acid |
| NMR | nuclear magnetic resonance |
| PGC | porous graphitized carbon chromatography |
| PMP | 1-phenyl-3methyl-5pyrazolone |
| PSD | post source decay |
| PTM | post-translational modifications |
| QbD | Quality by Design |
| QC | quality control |

| | |
|---|---|
| QTOF | quadrupole time-of-flight |
| RP | reversed-phase chromatography |
| sIgA | secretory immunoglobulin A |
| TFA | trifluoroacetic acid |
| TOF | time of flight |
| UDP-GalNAc | N-acetyl galactosaminyltransferase |
| UHPLC | ultra high performance liquid chromatography |
| UTI | urinary tract infection |
| WAX | weak anion exchange chromatography |

## Curriculum Vitae

Radoslaw Pawel Kozak was born on November 5th 1983, in Wloclawek, Poland. He attended high school in Plock, Poland from which he graduated in 2002. He then started his studies at University of Warsaw, Poland where he earned his M.Sc. degree in Organic Chemistry in June 2007. The research for his Masters thesis was performed in the Laboratory of Stereocontrolled Organic Synthesis under the supervision of Prof. Dr. R.R. Sicinski. The project focused on the synthesis of hydrindane precursors of vitamin D with modified side chains.

In 2008 he moved to Oxford, UK and in December 2008 he joined Ludger Ltd as a Scientist, in the Development and Glycoprofiling groups. Within those groups he has been involved in the development of new techniques for glycan release and derivitization for MS and HPLC/UPLC, and also glycan profiling of biopharmaceuticals. In June 2010 he began his PhD entitled: “Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins” in collaboration with Department of Parasitology of the Leiden University Medical Center. In January 2016 he was promoted to a Senior Scientist at Ludger Ltd.

## List of publications

[1] **R.P. Kozak**, L. Royle, R.A. Gardner, D.L. Fernandes, M. Wuhrer, Suppression of peeling during the release of O-glycans by hydrazinolysis. Anal.Biochem. 423 (2012) 119-128.

[2] G. Zauner, **R.P. Kozak**, R.A. Gardner, D.L. Fernandes, A.M. Deelder, M. Wuhrer, Protein O-glycosylation analysis. Biol.Chem. 393 (2012) 687-708.

[3] C. Gao, Y. Liu, H. Zhang, Y. Zhang, M.N. Fukuda, A.S. Palma, **R.P. Kozak**, R.A. Childs, M. Nonaka, Z. Li, D.L. Siegel, P. Hanfland, D.M. Peehl, W. Chai, M.I. Greene, T. Feizi, Carbohydrate sequence of the prostate cancer-associated antigen F77 assigned by a mucin O-glycome designer array. J.Biol.Chem. 289 (2014) 16462-16477.

[4] G.V. Harlalka, A. Lehman, B. Chioza, E.L. Baple, R. Maroofian, H. Cross, A. Sreekantan-Nair, D.A. Priestman, S. Al-Turki, M.E. McEntagart, C. Proukakis, L. Royle, **R.P. Kozak**, L. Bastaki, M. Patton, K. Wagner, R. Coblentz, J. Price, M. Mezei, K. Schlade-Bartusiak, F.M. Platt, M.E. Hurles, A.H. Crosby, Mutations in B4GALNT1 (GM2 synthase) underlie a new disorder of ganglioside biosynthesis. Brain 136 (2013) 3618-3624.

[5] **R.P. Kozak**, L. Royle, R.A. Gardner, A. Bondt, D.L. Fernandes, M. Wuhrer, Improved nonreductive O-glycan release by hydrazinolysis with ethylenediaminetetraacetic acid addition. Anal.Biochem. 453 (2014) 29-37.

[6] A. van Diepen, A.J. van der Plas, **R.P. Kozak**, L. Royle, D.W. Dunne, C.H. Hokke, Development of a Schistosoma mansoni shotgun O-glycan microarray and

application to the discovery of new antigenic schistosome glycan motifs. Int.J.Parasitol. 45 (2015) 465-475.

[7] **R.P. Kozak**, C.B. Tortosa, D.L. Fernandes, D.I. Spencer, Comparison of procainamide and 2-aminobenzamide labeling for profiling and identification of glycans by liquid chromatography with fluorescence detection coupled to electrospray ionization-mass spectrometry. Anal.Biochem. 486 (2015) 38-40.

[8] V. Dotz, R. Haselberg, A. Shubhakar, **R.P. Kozak**, D, Flack, Y. Rombouts, D. Reusch, G.W. Somsen, D.L. Fernandes, M. Wuhrer, Mass spectrometry for glycosylation analysis of biopharmaceuticals. Trends in Analytical Chemistry. 73 (2015) 1-9.

[9] C. Phansopa, **R.P. Kozak**, L.P. Liew, A.M. Frey, T. Farmilo, J.L. parker, D.J. Kelly, R.J. Emery, R.I. Thompson, L. Royle, R.A. Gardner, D.I. Spencer, G.P. Stafford,Characterization of a sialate O-acetyl esterase (NanS) from the oral pathogen Tannerella forsythia that enhances sialic acid release by NanH its cognate sialidase. Biochemical Journal. 472 (2015) 157-167.

[10] A. Bondt, S. Nicolardi, B.P. Jansen, K. Stavenhagen, D. Blank, G.S. Kammeijer, **R.P. Kozak**, D.L. Fernandes, P.J. Hensbergen, J.M. Hazes, Y.E. van der Burgt, R.J. Dolhain, M. Wuhrer, Longitudinal monitoring of immunoglobulin A glycosylation during pregnancy by simultaneous MALDI-FTICR-MS analysis of N- and O-glycopeptides. Scientific Reports. 6 (2016) 1-12.

[11] A. Shubhakar, **R.P. Kozak**, K.R. Reiding, L. Royle, D.I. Spencer, D.L. Fernandes, M. Wuhrer, Automated High-Throughput Permethylation for

Glycosylation Analysis of Biologics Using MALDI-TOF-MS. Anal Chem. 17 (2016) 8562-9.

[12] **R.P. Kozak**, P.A. Urbanowicz, C. Punyadeera, K.R. Reiding, B.C. Jansen, L. Royle, D.I. Spencer, D.L. Fernandes, M. Wuhrer, Variation of human salivary O-glycome. PLoS One. 9 (2016) 1-15.

## Acknowledgements

I would like to express my appreciation and thanks to my PhD advisor and CEO of Ludger Ltd. Dr Daryl Fernandes, for encouraging my research and for allowing me to start my PhD journey and grow as a glyco-scientist.

Special mention goes to my enthusiastic supervisor Professor Manfred Wuhrer for his patience and guidance. His knowledge of glycosylation and experience in the development of glycoanalytical techniques especially for mass spectrometry is of great value to PhD student. My entire PhD progressed with many constructive discussions with him on data analysis, writing manuscripts and my thesis.

Dozens of people have helped and taught me immensely at Ludger. Dr Richard Gardner welcomed me to the lab and introduced the hydrazinolysis procedure. Dr Daniel Spencer introduced mass spectrometry, shared his knowledge and time to explain the importance of mass spectrometry in the glyco-world.

I would like to take this opportunity also to thank Dr Louise Royle from Ludger. She is a good teacher and has offered great help and guidance during my PhD. She is very well organised and she always thinks ahead of time. I have received tremendously valuable advice from her on how to perform good experiments and how to organise my tasks in the most efficient way. She shared with me her great experience in experimental procedures, especially in O-glycan release, exoglycosidase sequencing and O-glycan analysis. I appreciate her comments and suggestions.

I would also like to thank all of my friends and colleagues from Ludger and LUMC who supported me in the lab, in writing, and encouraged me to strive towards my goal.

Lastly, I would like to thank my supportive, encouraging and patient wife Marta whose support during different stages of this PhD has been so appreciated. Thank you.