Glycan mapping of glycoproteins with UPLC-FLR-MALDI/TOF-MS

Sarah Bertilsson

Uppsala 2014-05-27

Supervisors: Ahmad Amini and Peter Ajdert, the Medical Product Agency and Margareta Ramström
Uppsala University

Subject specialist: Jonas Bergquist, Uppsala University

Examiner: Christer Elvingson, Uppsala University

UPPSALA UNIVERSITET

LÄKEMEDELSVERKET
MEDICAL PRODUCTS AGENCY
Abstract

Glycans, carbohydrate chains that are attached to proteins, have essential rolls in various biological processes, e.g., in folding and protein activity, but the glycosylation pattern on proteins changes with production conditions such as different pH. These patterns can be compared to a fingerprint and thus they can be used as a quality attribute. The object of the work presented in this thesis was to optimize the sample preparation and separation conditions in order to collect the glycans for identification and quantification. Several glycoproteins were deglycosylated with two different protocols together with the enzyme peptide-N-glycosidase F, which was precipitated with ethanol or acetone and two labels (anthranilic acid and anthranileamide) were compared and analysed with ultra performance liquid chromatography. The glycosylation protocol using a deglycosylation kit was more time effective than the other protocol. The anthranilic acid label was more sensitive than the other label for analysis of matrix-assisted laser desorption/ionisation. The method validation was carried out with glycans released from the glycoprotein fetuin. Method validation displayed a relative standard deviation of 0.1% for intra-day precision for retention time and 0.7% for peak area. Inter-day precision showed a relative standard deviation below 3.2 % for the retention times and below 34.2% for the peak areas. The performed significance test showed that the method could only be used for identification of glycans but not for relative quantitative measurements.
Abbreviations

2-AA – anthranilic acid
2-AB - 2-aminobenzamide
ACN - acetonitrile
Asn - asparagine
Asp - aspartic acid
CHCA- α-cyano-4-hydroxycinnamic acid
CHO – Chinese hamster ovary cell
DHB- dihydroxybenzoic acid
DMSO- dimethyl sulfoxide
DTT – dithiothreitol
EDTA- ethylenediaminetetraacetic acid
EPO – erythropoietin
ER – endoplasmic reticulum
FLR – fluorescence detector
Fuc - fucose
GalNAc – N-Acetylgalactosamine
Glc – glucose
GlcNAc- N-Acetylglucosamine
HETP – height equivalent to a theoretical plate
HILIC-hydrophilic interaction liquid chromatography
HPAEC-PAD – high performance anion-exchange chromatography with pulsed amperometric detection
HPLC – high performance liquid chromatography
IAA- iodoacetamide
LC –liquid chromatography
CE – capillary electrophoresis
MALDI/TOF MS- matrix assisted laser desorption/ionisation time-of-flight mass spectrometry
Man - mannose
MS – mass spectrometry
NEB- New England BioLabs®
Neu5Ac – N-Acetylneuraminic acid
NMR – nuclear magnetic resonance
PAS – periodic acid-Schiff
PNGase F - peptide-N-glycosidase F
PTM – post-translation modification
RSD – relative standard deviation
Ser – serine
Sia- sialic acid
TFA – trifluoroacetic acid
THAP- 2’,4’,6’-trihydroxyacetophenone monohydrate
Thr - threonine
Tris- tris(hydroxymethyl)aminomethane
UPLC- ultra performance liquid chromatography
UV- ultraviolet
Xyl - xylose
Acknowledgements

This thesis has been carried out at the Medical Product Agency (MPA; Läkemedelsverket) and for two years I have been studying analytical chemistry at Uppsala University where I have gained knowledge about various areas. I would like to express my sincere gratitude towards a number of people:

Ahmad Amini, my supervisor for this thesis. Thank you for all guidance during the project and the very interesting conversations during fika and lunch!

Peter Ajdert, my co-supervisor, without your knowledge about the instruments I would not have been able to perform any analysis at all!

All the staff at the MPA, thank you for the warm welcome and much needed help during the project.

All of the professors, lecturers, PhD’s and students during my two years at Uppsala thank you! I would not have known anything without you.

Anna, the world’s best study buddy, studying for the exams would have been horrible without you.

Minéa, my dearest friend, without you I would not have felt like 22 again!

My parents and my sister, without your support I do not known what I would have done. You’re the best!
### Table of contents

1. Introduction .................................................................................................................. 1
2. The chemistry of carbohydrates ................................................................................. 1
   2.1 Glycans ................................................................................................................... 2
   2.2 Glycosylation .......................................................................................................... 3
3. Analysis of glycans ....................................................................................................... 4
   3.1 Release of glycans and labelling ............................................................................. 4
   3.2 Ultra performance liquid chromatography ............................................................... 6
      3.2.1 Hydrophilic interaction liquid chromatography ................................................. 7
3.3 Detectors .................................................................................................................... 9
   3.3.1 Fluorescence detector ......................................................................................... 9
   3.3.2 Matrix-assisted laser desorption ionisation - Time-of-Flight mass spectrometer .... 9
4. Materials and Methods ............................................................................................... 11
   4.1 Proteins .................................................................................................................. 11
   4.2 Chemicals .............................................................................................................. 11
5. Instrumentation .......................................................................................................... 12
6. Experimental .............................................................................................................. 12
   6.1 General UPLC experimental .................................................................................. 13
   6.2 Deglycosylation protocol I .................................................................................... 13
   6.3 Deglycosylation with protocol I and II .................................................................. 15
   6.4 Precipitation with acetone or ethanol ................................................................. 15
   6.5 Labelling with 2-AA or 2-AB ............................................................................... 16
   6.6 Matrices for MALDI ............................................................................................. 16
   6.7 EPO and follitropin ............................................................................................... 17
   6.8 Validation of method for fetuin .............................................................................. 18
7. Results and discussion .............................................................................................. 18
   7.1 Deglycosylation of several proteins ..................................................................... 18
   7.2 Deglycosylation protocols .................................................................................... 19
   7.3 Precipitation with acetone or ethanol .................................................................... 23
   7.4 Labelling with 2-AA or 2-AB ................................................................................. 24
   7.5 Matrices for MALDI .............................................................................................. 26
   7.6 EPO and follitropin ............................................................................................... 28
   7.7 Method validation for fetuin .................................................................................. 32
8. Conclusion .................................................................................................................. 36
9. Future aspects ............................................................................................................. 36
10. Popular summary.................................................................................................................. 37
References .................................................................................................................................. 39
Appendix A .................................................................................................................................. 40
  Mass spectra from glycans labelled with 2-AB........................................................................ 40
  Mass spectra from glycans labelled with 2-AA........................................................................ 43
1. Introduction

Proteins can undergo a diversity of different post-translation modifications (PTMs). Among these, glycosylation is the most common and also the most complex modification which only occurs in eukaryotic cells. It has been estimated that more than 50% of all proteins undergo this type of modification [1, 2]. PTMs can be seen as a finale fine-tuning of the proteins that influence different properties or activate the proteins.

Glycosylation involves a successive covalent attachment of monosaccharides to the protein during the synthesis and passage through the endoplasmic reticulum (ER) and the Golgi apparatus. This ultimately creates a specific and complex glycosylation pattern that is important to characterize, since the glycans are involved in a variety of important processes in the body, such as folding, stability, cell-cell interaction, receptor activation, antigenicity, pharmacokinetics, pharmacodynamics, half-life and molecular trafficking [3-6]. Changes (e.g. overexpression, underexpression and disease-associated glycans) in the glycosylation pattern have essential impact on the three-dimensional structure of the proteins and consequently the anticipated and sought function is altered [7].

The glycosylation depends on production conditions. Changes in pH, presence of nutrients and hormones will yield different glycosylation patterns. Biopharmaceuticals manufactured by different companies will yield different glycosylation as they are expressed in different living systems (e.g. mice, yeast or Chinese hamster ovary cells (CHO)) [5, 8-10]. These changes in glycosylation affect many functions and in the end these can change the biological activity[11].

For this thesis, the Medical Products Agency wanted to develop a method for glycan mapping of various recombinant proteins to be able to investigate the quality and relative quantity of glycans. This could be used as a method for controlling biopharmaceuticals, e.g., batch-to-batch variations, degree of glycosylation or glycosylation patterns between biosimilars.

2. The chemistry of carbohydrates

In nature, there are hundreds of different monosaccharides but in higher animals they are limited to a few, which are shown in Table 1[6]. The monosaccharides can be assembled in different manners in the endoplasmic reticulum (ER) and the Golgi apparatus through a glycosidic bond between the hemiacetal group on the first monosaccharide and the hydroxyl
group on the second. Since the monosaccharides have several possible hydroxyl groups which can bind to the hemiacetal group, a branching structure can be created. Another layer of complexity is added as the hydroxyl group can have two different orientations, \( \alpha \) (axial) or \( \beta \) (equatorial) depending on stability[12]. This ultimately means that the glycans varies a lot in their structure.

<table>
<thead>
<tr>
<th>Subclass</th>
<th>Description</th>
<th>Monosaccharide</th>
<th>Abbreviations</th>
<th>( M_w ) (u)</th>
<th>Structure</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pentose</td>
<td>Five-carbon neutral sugar</td>
<td>Xylose</td>
<td>Xyl</td>
<td>132.04</td>
<td><img src="image" alt="Xylose" /></td>
<td></td>
</tr>
<tr>
<td>Hexoses</td>
<td>Six-carbon neutral sugar</td>
<td>D-Mannose</td>
<td>Man</td>
<td>162.05</td>
<td><img src="image" alt="D-Mannose" /></td>
<td><img src="image" alt="Green" /></td>
</tr>
<tr>
<td></td>
<td></td>
<td>D-Galactose</td>
<td>Gal</td>
<td>162.05</td>
<td><img src="image" alt="D-Galactose" /></td>
<td><img src="image" alt="Yellow" /></td>
</tr>
<tr>
<td></td>
<td></td>
<td>D-Glucose</td>
<td>Glc</td>
<td>162.05</td>
<td><img src="image" alt="D-Glucose" /></td>
<td><img src="image" alt="Blue" /></td>
</tr>
<tr>
<td>Deoxyhexose</td>
<td>Hex without the hydroxyl group at the 6-position</td>
<td>L-Fucose</td>
<td>Fuc</td>
<td>145.06</td>
<td><img src="image" alt="L-Fucose" /></td>
<td><img src="image" alt="Red" /></td>
</tr>
<tr>
<td>Hexosamines</td>
<td>Hex with an amino group or it's derivate at 2-position</td>
<td>N-Acetylgalactosamine</td>
<td>GalNAc</td>
<td>203.08</td>
<td><img src="image" alt="N-Acetylgalactosamine" /></td>
<td><img src="image" alt="Yellow" /></td>
</tr>
<tr>
<td></td>
<td></td>
<td>N-Acetylgucoseamine</td>
<td>GlcNAc</td>
<td>203.08</td>
<td><img src="image" alt="N-Acetylgucoseamine" /></td>
<td><img src="image" alt="Blue" /></td>
</tr>
<tr>
<td>Sialic Acid</td>
<td>Nine-carbon acidic sugar</td>
<td>e.g.</td>
<td>Sia</td>
<td>291.10</td>
<td><img src="image" alt="Sialic Acid" /></td>
<td><img src="image" alt="Purple" /></td>
</tr>
</tbody>
</table>

Table 1. Information about the classes, names, structure and symbols of the different monosaccharides in higher animals. All glycans illustrated in this thesis is done with program GlycoWorkbench[13].

2.1 Glycans

Glycans are divided into different classes depending on which amino acid they are bound to and the most common classes are N-linked and O-linked glycans. N-linked glycans are
attached to the nitrogen atom in asparagine (Asn) in the sequence Asn-X-Ser/Thr (serine/threonine), where X can be any amino acid except proline, and O-linked are attached an oxygen atom in the amino acids Ser or Thr[6, 12]. Of these two types, N-linked glycans are most studied and understood and this thesis will only investigate those. Below, glycans will thus mean N-linked glycans.

2.2 Glycosylation

As the process of glycosylation is complex and involves a series of synchronized enzyme reactions, it can be simplified by dividing it into three steps (Figure 1). The first two steps are formation of the precursor glycan followed by transfer of this precursor to the protein. These events occur in the different parts of the ER. The third step occurs in the different apartments of the Golgi apparatus in which the glycans is modified [6, 12]. The assembly of the precursor glycan starts in the cytoplasm of the ER with a series of enzyme reactions that link the dolichol molecule in the lipid bilayer to two GlcNAc and five mannose. During a mechanism that is unknown the dolichol-monosaccharide complex flips over to the lumen of the ER. In the lumen, four mannose are attached to the glycan together with three glucose. The second part follows after the transfer of the precursor glycan to the unfolded protein through an enzyme reaction. The third part starts in the ER with the removal of the three glucose, which indicates that the glycosylation is done and the glycoprotein is ready for modification in the Golgi apparatus. The glycoprotein is secreted to the Golgi and the modifications start with the removal of some mannose.

The N-linked glycans have a common core structure consisting of two GlcNAc and three mannose. The glycans can be divided into three types: high mannose, complex and hybrid (Figure 1). The only modification the high mannose glycan will receive is the removal of the mannose directly in the Golgi. Complex glycans are re-elongated and modified with the help of several enzymes and they can have several antennas. Hybrid glycans are generated when a certain enzyme reaction does not occur and the consequence is that some mannose are not removed, but the re-elongation continues [6, 12].
3. Analysis of glycans

In order to analyze the glycans they have to be cleaved from the glycoprotein, which is usually done with the help of specific enzymes. A complication encountered when analyzing glycans is the lack of UV-absorbing properties. There are techniques such as high-pH anion-exchange chromatography-pulsed amperometric detection (HPAEC-PAD) and periodic acid-Schiff reaction (PAS) in which the lack of chromophore does not interfere with detection. Other techniques such as liquid chromatography (LC) and capillary electrophoresis (CE) are useful if the glycans are labelled with a chromophore before analysis. If these techniques are coupled to a mass spectrometer (MS) they are even more powerful. Nuclear magnetic resonance (NMR) is an excellent technique for determination of structures and the stereochemistry but this technique calls for much higher concentration compared to what is normally achieved [12].

3.1 Release of glycans and labelling

The most common way of releasing glycans is through an enzymatic reaction, but another possibility is a chemical release with hydrazine which is explosive and should not be handled.
daily. There are various enzymes available with different specificity; this leads to an arsenal of possible bonds to break. Digestion with the enzyme peptide-N-glycosidase F (PNGase F) is a popular approach when analyzing glycans. PNGase F hydrolyses the bond between N and C in the amino acid Asn resulting in a reduced end for the glycan and the Asn is converted to Asp (Figure 2) [6, 12]. PNGase F is not able to release glycans which have a α1-3-linked fucose at the innermost monosaccharide, which is a feature that can be seen in plants and insects and thus not a common problem [14].

*Figure 2. The enzyme PNGase F selectively separates the glycans linked to Asn in the protein by hydrolysis of the N-C bond and converts Asn to the amino acid Asp.*

After the release, the glycans are labelled with a fluorophore through a reductive amination reaction (Figure 3). There are several fluorophores which have been used for LC-analysis, e.g., 2-aminobenzamide (anthranilamide, 2-AB) or O-aminobenzoic acid (anthranilic acid, 2-AA). Lately it has been reported that the 2-AA label yielded both neutral and acidic glycans in the same mass spectrum in negative mode MALDI/TOF-MS [3]. This solves a problem encountered when analyzing glycans with MALDI/TOF-MS as the negative glycans (sialic acids) are unstable and can degrade before reaching the detector.
Figure 3. The reductive amination reaction used for labelling of the glycans. The glycan is labelled by formation of a Schiff’s base between the glycan and label. The driving force of the reaction is the reduction of the double bond between the carbon and the nitrogen.

3.2 Ultra performance liquid chromatography

Liquid chromatography has been used for over 30 years and high performance liquid chromatography (HPLC) is a common technique for analyzing a broad range of substances. A sample is injected into column that is filled with a stationary phase, and a mobile phase is pumped through the column, which carries the analytes through the system. The analytes in the sample will be separated due to the degree of distribution between the mobile and stationary phase [15].

The van Deemter equation (Equation 1) describes the three kinetic processes that influence the band broadening in the chromatographic separation. The term A represents Eddy diffusion where the band broadening of the peaks is explained by the analytes’ possibility to travel different ways through the column. This term is independent of flow rate (u) and depends on the packing of the column [15].

The longitudinal diffusion is described by the B-term and says that the band broadening occurs in every direction due to concentration differences at the edges of the injected sample. This effect will increase as flow rate decreases. A high flow rate will thus reduce this effect[15].

As the analytes travels through the column, some will diffuse into the stationary phase and some will interact less with the stationary phase hence band broadening arises and this phenomenon is called mass transport (C). This term is highly dependent on the particle size and the flow rate. Small particles give less band broadening than larger particles as the analytes cannot penetrate the small particles to the same extent. Higher flow rate will increase band broadening accordingly to equation 1 [15].
\[ HETP = A + \frac{B}{u} + Cu \]

Eq. (1)

The demand for smaller particle sizes of the stationary phase creates a pressure problem and the conventional instruments can not handle pressures originating from smaller particle sizes than 3 µm [15].

In 2004 ultra performance liquid chromatography (UPLC) was introduced and has since then been more and more applied for different analyses. The UPLC-instrumentation can handle pressures from a more tightly packed sub-2 µm column and a smaller inner column diameter [16]. The smaller particle sizes will lead to faster separation and higher theoretical plate height/peak capacity [15].

Figure 4 shows the van Deemter equation curve, which displays the relationship between flow rate and HETP (height equivalent to a theoretical plate). According to the van Deemter curve a certain particle size has an optimum flow rate [16]. By looking at the curve for \( d_p = 1.7 \) µm in Figure 4 it is obvious that the curve has a larger optimum flow rate range than larger particles.

In the end a UPLC analysis will have higher resolution and shorter analysis compared with HPLC.

![Figure 4. The van Deemter curve shows that for smaller particle sizes the plate height is almost independent flow rate. Published with permission from the publisher[17].](image)

### 3.2.1 Hydrophilic interaction liquid chromatography

In 1990 Alpert introduced the abbreviation HILIC for hydrophilic interaction liquid chromatography as a variant of normal phase chromatography (a hydrophobic mobile phase
and hydrophilic column for eluting polar compounds, i.e., glycans) [18]. The main difference between normal phase and HILIC is the composition of the mobile phase. In normal phase the mobile phase consist of only an organic solvent that is immiscible with water. In HILIC the mobile phase is water or buffer and a water-miscible organic solvent. This will lead to different separation mechanisms, partitioning for HILIC and surface adsorption for normal phase[19].

The partitioning mechanism is complex in HILIC, but it is believed that a water layer will form in the stationary phase making the stationary phase more polar (Figure 5)[19]. This distinct water/mobile phase difference separates the analytes in the same manner as in liquid-liquid extraction. Another dimension of retention is added as the analytes interact with the stationary phase through hydrogen bonding, dipole-dipole and electrostatic interaction. This can be controlled by the selection of the stationary phase[20]. The different interactions are illustrated in Figure 5.

As not all laboratories have access to an UPLC instrument equipped with a HILIC-column there exists a possibility to separate the glycans with a HPLC equipped with column with halo particles or solid core particles [21].

Figure 5. Hydrophilic partitioning, hydrogen bonding and electrostatic interaction are the three of the four different types of interaction that influences the retention time of the glycans. Dipole-Dipole interactions are not seen here. Published with the permission from the author[22].
3.3 Detectors

3.3.1 Fluorescence detector
This detector is based on the ability of a substance to fluoresce and it is a very sensitive detector, up to 100-1000 times more sensitive than a UV-detector[23]. The analyte is excited by a wavelength close to the absorbance maximum for UV-light, in this excited state the analyte is relaxed to the lowest energy level of the excited state by collisions with other analytes and heat is released. When the analyte is relaxed from the excited state to the ground state, light is emitted at a lower energy (higher λ) compared the wavelength for the excitation since some energy was lost as heat [24].

3.3.2 Matrix-assisted laser desorption ionisation- Time-of-Flight mass spectrometer
Mass spectrometry (MS) is a technique which measures the mass-to-charge ratio (m/z) of unknown compounds and uses the m/z for identification. A general scheme for a mass spectrometer is shown in Figure 6. The sample is introduced to an ion source which ionises the analytes and transfer them to the mass analyser. The generated ions will be separated in the mass analyser after their m/z and the detector will measure the intensity of the ions and convert it an electric signal [25].

Figure 6. A schematic overview of the different parts of a mass spectrometer. The sample is introduced and ionised in the ion source followed by separation of the analytes in the mass analyser and the signal is converted to an electric signal in the detector.

Matrix-assisted laser desorption ionisation- time-of-flight (MALDI/TOF-MS) is a mass spectrometer consisting of two parts, MALDI and TOF. MALDI is the ion source and is a soft ionisation technique in which the sample is mixed with a matrix, applied on a target, inserted and irradiated with laser (Figure 7A). Upon irradiation, the analytes will be ionised in the gas
phase with the help of the matrix through a mechanism which is not fully understood. This technique is especially suitable for molecules that are large, non-volatile and/or thermally labile. The ionisation can also be a drawback as the analytes can be ionised with different efficiency. Another drawback is the fact that the sample is crystallised heterogeneously on the target [25, 26].

There are several possible matrices available and the choice is made depending on the analyte together with the desired amount of energy to be absorbed. Common matrices such as 2,5-dihydroxybenzoic acid (DHB) and α-cyano-4-hydroxycinnamic acid (CHCA) are known as “hotter” matrices as they absorb more energy. “Colder” matrices do not absorb the same amount of energy and 2’,4’,6’-trihydroxyacetophenone monohydrate(THAP) is categorized as a “colder” matrix [26].

Since the laser is pulsed on the sample spot, a perfect mass analyser coupled to a MALDI source is a TOF (Figure 7B). A short time (100-500 ns) after the laser pulse, a potential (positive or negative) is applied creating an electric field which accelerates the ions into the field-free tube [26]. All of the ions get almost the same kinetic energy and will enter the tube with different velocities depending on their masses. Equation 2 shows that the velocity, \( v \), depends on the total charge \( q = ze \), the applied potential \( V_s \) and the mass \( m \). The time \( t \) that the ion will spend in the tube is calculated according to Equation 3, where \( L \) is the length of the tube. These two equations are combined into equation 4 which shows that \( m/z \) can be estimated from the time spent in the tube [25].

\[
\begin{align*}
    v &= \sqrt{\frac{2zeV_s}{m}} \quad \text{Eq. (2)} \\
    t &= \frac{L}{v} \quad \text{Eq. (3)} \\
    t^2 &= \frac{m}{z} \left( \frac{L^2}{2eV_s} \right) \quad \text{Eq. (4)}
\end{align*}
\]

There are two modes of TOF, linear and reflectron. The reflectron (ion mirror) improves the mass resolution by creating an electrostatic field which corrects the differences in the velocity for ions with the same \( m/z \), but the sensitivity is decreased [25].
Figure 7. The principles of MALDI and TOF A) The ionisation with MALDI. The sample/matrix mixture is irradiated with a laser and the analyte will be ionised in gas phase. B) The ions are inserted into the TOF and accelerated through a field free area. The ions will be separated because the velocity is dependent of the mass. Published with permission from publisher [26].

4. Materials and Methods

4.1 Proteins

Different proteins were used for the experiments. Ribonuclease B, ovalbumin and fetuin were purchased from Sigma (St. Louis, MO, USA), IgG and follitropin were reference standards from European Pharmacopoeia, and EPREX® (EPO) was acquired from a drug store in Uppsala.

4.2 Chemicals

All of the chemicals were of pro analysis (p.a) grade and the water used was of MilliQ quality. A deglycosylation kit from New England Biolabs® (NEB) were used to deglycosylate the proteins. Acetonitrile, methanol, ethanol, acetone, ethylenediaminetetraacetic acid (EDTA), urea, ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide (IAA), NaBH₃CN, 2-AB, 2-AA and 2-methyl-pyridine borane complex were purchased from Sigma (St. Louis, MO, USA). Dimethyl sulfoxide (DMSO), glacial acetic acid, tris(hydroxymethyl)aminomethane (Tris), ortho-phosphoric acid (85%), formic acid (98-100%) and ammonium acetate were bought from Merck (Darmstadt, Germany). Ammonium formate was purchased from Alfa Aesar (Ward Hill, MA, USA).
**5. Instrumentation**

For all the reactions where heating was required an Eppendorf Thermomixer® comfort from Eppendorf AG (Hamburg, Germany) was used.

For all sample preparation involving speedvac, a DNA120 SpeedVac® system from Thermoelectron corporation (Milford, MA, US) was used.

For the ultrasonic baths a Bandelin SONOREX™ Digital 10 P ultrasonic bath from Sigma (St. Louis, MO, US) was used.

Centrifugation was performed with a microcentrifuge 4214 from ALC (Kontron, UK).

For the solid phase extraction, GlycoWorks™ HILIC 1 cc 10 mg (Waters, Milford, MA, US) columns were used.

An ACQUITY UPLC® system equipped with a 2.1x150 mm BEH Glycan 1.7 µm column (Waters, Milford, MA, US) was used for the chromatographic separations. An ACQUITY FLR detector (Waters, Milford, MA, US) was used at different excitation wavelengths depending on the label.

All of the mass spectrometric measurements were carried out on an Ultraflex II MALDI TOF MS (Bruker Daltonics, Bremen, Germany) with a pulsed nitrogen laser at 337 nm and the data analysis were done in the software Autoflex. The analysis was performed at linear positive or negative mode. The instruments were calibrated before every measurement with Bruker Daltonics standard peptide mixture consisting of six peptides (angiotensin I, angiotensin II, substance P, bombesin, ACTH clip 1-17, ACTH clip 18-39).

**6. Experimental**

In Figure 8 a schematic overview of the workflow for the experiments is shown. First different denaturation and deglycosylation protocols were investigated with ribonuclease B, ovalbumin, IgG and fetuin in order to find the best deglycosylation conditions. After this the glycoprotein fetuin was chosen for further experiments as fetuin contains sialic acids. Two different solvents were investigated for the precipitation of the protein. After this two different labels were tested in order to find the best one for MALDI/TOF-MS and UPLC. The developed method was then validated. Two different recombinant proteins, EPO from EPREX® and follitropin were then deglycosylated and analysed with different gradients in UPLC to find the best separation conditions.
Figure 8. A schematic overview of the analysis of glycans in this thesis. First of all, the protein is denatured followed by addition of the enzyme that cleaves the glycans from the protein. The glycan is then labelled and analysed with UPLC and MALDI/TOF-MS.

6.1 General UPLC experimental

For the UPLC analyses, the column temperature was 60 °C and 2 µL sample was injected into the column. The FLR was set at $\lambda_{ex} = 260$ nm / $\lambda_{em} = 420$ nm for the 2-AB labelled glycans and $\lambda_{ex} = 360$ nm / $\lambda_{em} = 430$ nm for the 2-AA labelled glycans.

For the mobile phase A (100 mM ammonium formate) a stock solution of 1 M was prepared. The mobile phase A was prepared by taking 100 mL from the 1 M solution and the pH was adjusted to 4.5 with diluted formic acid and diluted to 1000 mL. This solution was filtered through a 0.22 µm filter and placed in an ultrasonic bath for 30 minutes. The mobile phase B was acetonitrile.

6.2 Deglycosylation protocol I

The proteins ribonuclease B, ovalbumin, IgG and fetuin were deglycosylated according to the following protocol:

1. Protein solutions of 10 mg/mL were made by adding 10 mg of protein to 1 mL of MilliQ water.
2. Combining 800 µg of respective protein with 500 µL denaturing buffer (8 M urea, 3 mM EDTA, 100 mM Tris and pH 9) and 3 µL of 45 mM DTT after which the mixture was incubated at 50 °C for 30 minutes.
3. After the samples had cooled down to room temperature, 5 µL of 100 mM IAA was added and the samples were stored at 4°C for 1 hour.
4. The denatured samples were desalted by Amicon® Ultra 3K centrifugal filter (Merck Millipore, Darmstadt, Germany), washed with 500 µL MilliQ water and centrifuged
down into 50 µL of 50 mM ammonium bicarbonate buffer (pH 7.9) and 2 µL PNGase F was added

5. All of the samples were incubated at 37 °C over night.

6. In the morning 500 µL ice-cold acetone was added to the samples, the samples were stored at -20 °C for 1 hour, and centrifuged at 14 000 rpm for 10 minutes. The supernatants were transferred to new tubes and the pellets were extracted with 3x30 µL ice-cold 60% methanol. All of the samples were placed in the speed vac until dry.

At the same time undenatured samples for each protein were prepared by:

1. Combining 100 µg of each protein with 50 µL of 50 mM ammonium bicarbonate buffer (pH 7.9) together with 2 µL of PNGase F.
2. All of the samples were incubated at 37 °C over night.
3. As point 6 above.

The labelling was done according to Bigge et al. by preparing a solution with concentration 1M of NaBH₃ and 0.35 M of 2-AB in DMSO/glacial acetic acid (70%/30% v/v) and adding 30 µL to each sample and incubated it at 65 °C for 2 hours [27]. Excess label was taken away by purification with solid phase extraction (SPE) columns from GlycoWorks™ by gravitation. The protocol followed the supplied instruction and can briefly be described by:

1. Wet the column with 1000 µL MilliQ water
2. Condition the column with 1000 µL 85 %ACN
3. Mix the sample with the amount of ACN to reach 85% ACN
4. Load on column
5. Wash with 3x400 µL 85 % ACN
6. Elute the glycans with 3x100 µL 100 mM ammonium acetate.

The eluate was analysed by UPLC with the gradient in Table 2 and the peaks were identified with the help of literature by matching chromatograms.
Table 2. The gradient used for analysing denatured and undenatured deglycosylated ribonuclease B, ovalbumin, IgG and fetuin.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Mobile phase A (%)</th>
<th>Mobile phase B (%)</th>
<th>Flow rate (ml/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-46.5</td>
<td>25 -&gt; 40</td>
<td>75 -&gt; 60</td>
<td>0.5</td>
</tr>
<tr>
<td>46.5-48</td>
<td>40 -&gt; 100</td>
<td>60 -&gt; 0</td>
<td>0.5</td>
</tr>
<tr>
<td>48-49</td>
<td>100</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>49-63</td>
<td>100 -&gt; 25</td>
<td>0 -&gt; 75</td>
<td>0.5</td>
</tr>
</tbody>
</table>

6.3 Deglycosylation with protocol I and II

Two different protocols (I and II) for denaturing and deglycosylation were investigated with ribonuclease B. Protocol I was the same except that the buffer was 100 mM and 50 mM phosphate buffer instead of the 50 mM (pH 7.9) ammonium bicarbonate buffer in point 4, section 5.2.

Protocol II followed the instruction from a deglycosylation kit manufactured by New England Biolabs®:

1. 20 μg glycoprotein was mixed with 1 μL 10X Glycoprotein Denaturing Buffert (400 mM DTT and 5% SDS) and MilliQ water for a final volume of 10 μL. This mixture was allowed to react for 10 minutes at 100 °C.
2. When the sample had cooled down, 2 μL of 10X G7 Reaction Buffer (500 mM sodium phosphate, pH 7.5), 2 μL of 10% NP-40 (detergent), MilliQ water and 2 μL PNGase F was added to reach a final volume of 20 μl.

The samples were labelled as in protocol I and analyzed with UPLC with the gradient in Table 2.

6.4 Precipitation with acetone or ethanol

For this investigation 40 μg fetuin was deglycosylated with the deglycosylation kit as in protocol II, but before the precipitation, the sample was divided into two aliquots. In one, the protein was precipitated with ice-cold acetone and the other was precipitated with ice-cold ethanol and both samples were stored in -20 °C for 1 hour before centrifugation for 10 minutes. The resulting supernatants were transferred to new tubes and the pellets were extracted with 3x30 μl ice-cold methanol and collected [28]. Both the supernatants and the
extracts from pellets were placed in a speed vac until dry. The samples were labelled with 2-AB as earlier and excess label was removed with SPE as earlier and injected in the UPLC with the gradient in Table 3.

Table 3. The gradient used for the analysis of the glycans from fetuin, where the protein had been precipitated with acetone or ethanol.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Mobile phase A (%)</th>
<th>Mobile phase B (%)</th>
<th>Flow rate (ml/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-100</td>
<td>25 -&gt; 40</td>
<td>75 -&gt; 60</td>
<td>0.5</td>
</tr>
<tr>
<td>100-104</td>
<td>40 -&gt; 100</td>
<td>60 -&gt; 0</td>
<td>0.5</td>
</tr>
<tr>
<td>104-110</td>
<td>100</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>110-120</td>
<td>100 -&gt;25</td>
<td>0 -&gt; 75</td>
<td>0.5</td>
</tr>
</tbody>
</table>

6.5 Labelling with 2-AA or 2-AB

Two samples of fetuin (20 µg) were deglycosylated according to protocol II, precipitated with ethanol and the pellets were extracted with 60 % ice-cold methanol and combined with the supernatant and placed in the SpeedVac until dry. Two different solutions for the labelling were prepared:

1. 2-AB solution: 1 M NaBH₃CN and 0.35 M 2-AB in DMSO/glacial acetic acid (70%/30% v/v)
2. 2-AA solution: 1 M 2-picoline-borane and 0.35 M 2-AA in DMSO/glacial acetic acid (85%/15% v/v)

To each sample 30 µL of the correct label was added and the samples were incubated at 65°C for 2 h. The excess label was removed as in protocol 1. The samples were placed in a speed vac until dry and injected into the UPLC with the gradient as in Table 2.

6.6 Matrices for MALDI

Two samples of fetuin (20 µg) were deglycosylated according to protocol II and labelled with 2-AB and 2-AA as in protocol I. After purification the samples were analysed with MALDI using six matrices (DHB, CHCA, THAP, THAP/DHB, THAP/CHCA and CHCA/DHB) and each matrix were analysed in positive and negative linear mode. The mixture of matrices was done by mixing 50%/50% (v/v) of each matrix in an Eppendorf tube. The DHB, CHCA
and THAP matrices were prepared by adding DHB or THAP into 50% ACN/50% of 0.2 % trifluoroacetic acid (TFA) in MilliQ until the solution was saturated.

6.7 EPO and follitropin

One syringe of 10 000 units of EPREX® containing the glycoprotein EPO was concentrated down to 0.43 µg/µL. The acquired follitropin solution had a concentration of 0.54 µg/µL. The proteins were deglycosylated (protocol II) and labelled with 2-AB as earlier and analysed with the UPLC but with different gradients. Follitropin were analysed with four different gradients (Tables 4, 5, 6 and 7). EPO was analysed with three different gradients, (Tables 2, 6 and 7).

Table 4. One of the gradients for follitropin.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Mobile phase A (%)</th>
<th>Mobile phase B (%)</th>
<th>Flow rate (ml/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-46.5</td>
<td>25 -&gt; 50</td>
<td>75 -&gt; 50</td>
<td>0.5</td>
</tr>
<tr>
<td>46.5-48</td>
<td>50 -&gt; 100</td>
<td>50 -&gt; 0</td>
<td>0.5</td>
</tr>
<tr>
<td>48-49</td>
<td>100</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>49-63</td>
<td>100 -&gt; 25</td>
<td>0 -&gt; 75</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 5. One of the gradients for follitropin.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Mobile phase A (%)</th>
<th>Mobile phase B (%)</th>
<th>Flow rate (ml/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-46.5</td>
<td>25 -&gt; 45</td>
<td>75 -&gt; 55</td>
<td>0.5</td>
</tr>
<tr>
<td>46.5-48</td>
<td>45 -&gt; 100</td>
<td>55 -&gt; 0</td>
<td>0.5</td>
</tr>
<tr>
<td>48-49</td>
<td>100</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>49-63</td>
<td>100 -&gt; 25</td>
<td>0 -&gt; 75</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 6. One of the gradients for both follitropin and EPO

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Mobile phase A (%)</th>
<th>Mobile phase B (%)</th>
<th>Flow rate (ml/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-80</td>
<td>25 -&gt; 45</td>
<td>75 -&gt; 55</td>
<td>0.5</td>
</tr>
<tr>
<td>80-84</td>
<td>45 -&gt; 100</td>
<td>55 -&gt; 0</td>
<td>0.5</td>
</tr>
<tr>
<td>84-85</td>
<td>100</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>86-100</td>
<td>100 -&gt; 25</td>
<td>0 -&gt; 75</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Table 7. One of the gradients for both follitropin and EPO.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Mobile phase A (%)</th>
<th>Mobile phase B (%)</th>
<th>Flow rate (ml/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-10</td>
<td>15 -&gt; 20</td>
<td>85 -&gt; 80</td>
<td>0.5</td>
</tr>
<tr>
<td>10-60</td>
<td>20 -&gt; 36</td>
<td>80 -&gt; 64</td>
<td>0.5</td>
</tr>
<tr>
<td>60-65</td>
<td>36 -&gt; 50</td>
<td>64 -&gt; 50</td>
<td>0.5</td>
</tr>
<tr>
<td>65-67</td>
<td>50</td>
<td>50</td>
<td>0.25</td>
</tr>
<tr>
<td>67-68</td>
<td>50-&gt;100</td>
<td>50-&gt;0</td>
<td>0.25</td>
</tr>
<tr>
<td>68-73</td>
<td>100</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>73-75</td>
<td>100-&gt;15</td>
<td>0-&gt;85</td>
<td>0.25</td>
</tr>
<tr>
<td>75-90</td>
<td>15</td>
<td>85</td>
<td>0.5</td>
</tr>
</tbody>
</table>

6.8 Validation of method for fetuin

In order to validate the method for fetuin, two samples (sample 1 and 2) of 60 µg fetuin were deglycosylated according protocol II, precipitated with ethanol, labelled with 2-AA and analysed with UPLC and the gradient in Table 2 but at different days. Three injections of each sample were done day 1 and day 2 in order to calculate the inter-day and intra-day precisions according to the Q2 validation guidelines from the International Conference on Harmonization (ICH) [29].

The robustness of the method was evaluated by deglycosylation of 60 µg fetuin (sample 3) as in protocol II but the temperature of the denaturation was changed from 100 °C to 95°C, the incubation time with PNGase F was changed to 22 hours and the labelling reaction temperature was decreased to 60 °C.

The relative standard deviation (RSD) for both retention time and ratio of peak area for two peaks were calculated for sample 1. For samples 1 and 2, the significance test t-test and f-test were performed in order to investigate if there was any difference between sample 1 and 2. The robustness was evaluated by a t-test comparing samples 1 and 3.

7. Results and discussion

7.1 Deglycosylation of several proteins

The deglycosylation of the ribonuclease B, ovalbumin, fetuin and IgG demonstrated that the undenatured glycoproteins did not display any glycan in the UPLC-analysis (data not
showed). This is not surprising as the folding often protects the glycans, thus the enzyme cannot cleave the glycans. Peaks corresponding to glycans were detected for all of the denatured samples (Figure 9a-d) and the glycans were assigned to the right peak by comparison with literature from UPLC analysis at the same conditions.

7.2 Deglycosylation protocols

Table 8 and Figures 10a-c show the resulting chromatograms and their data from the different deglycosylation protocols. These results clearly show that there are no differences between the different buffers or protocols concerning retention time for the different glycans (mannose 5 to mannose 9). The two chromatograms from the different phosphate buffers have smaller peaks (marked by red circles) before the Man5, Man6, Man8 and Man9 peaks that are not visible in the chromatogram for the deglycosylation kit. There is thus a possibility that there are components in the buffer that fluoresce at these wavelengths or that they are glycan fragments. Luckily they are well resolved from the main peaks and are thus not a problem. Another interesting observation is the relative concentration of glycans in the two phosphate buffers, i.e., the peak area of glycans in the 100 mM phosphate buffer is lower than the peak area for the 50 mM phosphate buffer (Table 8). This indicates that higher ionic strength decreases the enzyme activity. The deglycosylation with the kit is preferred as the sample preparation time is shorter and the buffer concentration for the reactions is the same as in the 50 mM phosphate buffer.
Table 8. The chromatographic data from Figure 10a-c including retention time and peak area.

<table>
<thead>
<tr>
<th></th>
<th>50 mM phosphate buffer</th>
<th>100mM phosphate buffer</th>
<th>Deglycosylation kit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Peak area (%)</td>
<td>Peak area (µV*sec)</td>
<td>Peak area (%)</td>
</tr>
<tr>
<td>Man5</td>
<td>48</td>
<td>37746570</td>
<td>44</td>
</tr>
<tr>
<td>Man6</td>
<td>36</td>
<td>28627932</td>
<td>38</td>
</tr>
<tr>
<td>Man7</td>
<td>3</td>
<td>2401159</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2755800</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>7430690</td>
<td>10</td>
</tr>
<tr>
<td>Man8</td>
<td>26</td>
<td>20242664</td>
<td>29</td>
</tr>
<tr>
<td>Man9</td>
<td>11</td>
<td>9024170</td>
<td>14</td>
</tr>
</tbody>
</table>
Figure 9. The chromatograms from the analysis of denatured a) IgG b) ribonuclease B c) fetuin and d) ovalbumin. The glycans were assigned to corresponding peaks with the help of literature.
Figure 10. The chromatograms from the different deglycosylation protocols for ribonuclease B.: a) 50 mM phosphate buffer, b) 100 mM phosphate buffer and c) the NEB kit. The red circle marks peaks that are not visible in the chromatogram for the NEB kit.
7.3 Precipitation with acetone or ethanol

When precipitating fetuin with the different solvents, the most striking difference between ethanol and acetone was found to be their efficiency to extract the glycan (Figures 11 and 12 and Table 9). Acetone was more efficient than ethanol as only 0.3 to 6 % of the glycans are in the extracts from the pellet (Table 9). For precipitation with ethanol the same percentage is 57- 87%. Inspection of the intensities of the chromatograms revealed that the concentrations of glycans are higher in ethanol than acetone. This could have been caused by an unequal aliquotation of glycans in the sample or that ethanol extracts the glycans better. In further sample preparation the proteins are precipitated with ethanol but the pellets are extracted with 60 % methanol and combined with the supernatant.

The relative peak area for the second group of peaks changes depending on the solvent. In ethanol, the first peak (t_R= 69.6 min) has the largest peak area followed by the second, third and fourth peak. For acetone, the most intense peak is the third (t_R= 76.4 min), followed by the third, second and lastly the first peak. This distribution indicates that larger glycans shows lower solubility in ethanol compared to acetone.

![Figure 11. Chromatograms for the glycans where the protein was precipitated with ethanol.](image1)

![Figure 12. Chromatograms for the glycans where the protein was precipitated with acetone.](image2)
Table 9. The area for the different peaks in Figures 11 and 12. With these values the percentage of glycans in the pellet were calculated.

<table>
<thead>
<tr>
<th>Peaks</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area for pellet extract</td>
<td>7979</td>
<td>769</td>
<td>1954</td>
<td>49291</td>
<td>28669</td>
<td>22794</td>
<td>4229</td>
</tr>
<tr>
<td>Area for supernatant</td>
<td>120923</td>
<td>55818</td>
<td>70089</td>
<td>1919571</td>
<td>1207066</td>
<td>3350026</td>
<td>1357153</td>
</tr>
<tr>
<td>% in pellet</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Peaks</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area for pellet extract</td>
<td>2214379</td>
<td>1519038</td>
<td>2159441</td>
<td>14078814</td>
<td>9031531</td>
<td>6238095</td>
<td>2188763</td>
</tr>
<tr>
<td>Area for supernatant</td>
<td>1642158</td>
<td>224318</td>
<td>585097</td>
<td>7368776</td>
<td>1590337</td>
<td>1524107</td>
<td>371597</td>
</tr>
<tr>
<td>% in pellet</td>
<td>57</td>
<td>87</td>
<td>79</td>
<td>66</td>
<td>85</td>
<td>80</td>
<td>85</td>
</tr>
</tbody>
</table>

7.4 Labelling with 2-AA or 2-AB

The labelling with 2-AA and 2-picoline-borane works well compared to 2-AB with NaBH$_3$CN. Table 10 and Figure 13 show that the more hydrophilic 2-AA gives longer retention times compared to the less hydrophilic 2-AB. The difference in retention times between the 2-AA and 2-AB peaks decrease, from 2.1 to 1.4 min, as the hydrophobicity of the glycan decreases (Chart 1). This demonstrates that the influence on the partitioning from the more hydrophilic 2-AA label is overruled by the larger size of the glycan, e.g., a more hydrophilic glycan.

The greatest advantages with 2-AA is the possibility to analyse both neutral and acidic (sialic acids) glycans in the same mass spectrum and that the reaction can be performed in aqueous conditions with the non-toxic 2-picoline-borance as the reductant [3].
Figure 13. The chromatograms resulting from the glycans released from fetuin and labelled with the two different labels. The top chromatogram belongs to the glycans labelled with 2-AA and has slightly longer retention times than the glycans labelled with 2-AB.

Table 10. The retention times for the different labeled peaks and the differences between them.

<table>
<thead>
<tr>
<th>2-AA</th>
<th>2-AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_R$ (min)</td>
<td>$t_R$ (min)</td>
</tr>
<tr>
<td>33.0</td>
<td>30.9</td>
</tr>
<tr>
<td>34.8</td>
<td>32.7</td>
</tr>
<tr>
<td>37.7</td>
<td>35.8</td>
</tr>
<tr>
<td>39.5</td>
<td>37.8</td>
</tr>
<tr>
<td>41.1</td>
<td>39.4</td>
</tr>
<tr>
<td>42.6</td>
<td>41.0</td>
</tr>
<tr>
<td>44.0</td>
<td>42.5</td>
</tr>
<tr>
<td>45.4</td>
<td>44.0</td>
</tr>
</tbody>
</table>

Chart 1. The difference between the 2-AB and 2-AA labelled peaks decreases as the hydrophobicity decreases demonstrating that the more hydrophilic a glycan is the less the label influences the separation.
7.5 Matrices for MALDI

When a MALDI-analysis is performed, it can be difficult to get a good spectrum as the analyte is not homogeneously distributed over the sample spot. During the analysis, a so-called “sweet spot” (a spot were the concentration of analytes is high) is desirable and the laser intensity is changed in order to hopefully retrieve a well-resolved spectrum. For the 2-AA labelled glycans the matrices THAP and DHB/THAP gave the most sensitive spectra in both positive and negative mode (Figures 14 and 16 and Appendix A). The matrix CHCA/DHB gave a sensitive spectrum in negative mode but not in positive mode (Figure 15). The m/z difference of 162 between the peaks in Figure 14 proves that a hexose has been cleaved off, in Figures 15 and 16 the difference is not 162 exactly but as the peaks are average masses, the peaks are broader than the monoisotopic peak giving an uncertainty in m/z. The 2-AB labelled glycans were more difficult to get a sensitive spectrum from. DHB and DHB/THAP (Figure 17 and 18) turned out to be the best matrices and these still showed a much lower sensitivity compared to the 2-AA labelled glycans. The spectra that are not presented here can be found in Appendix A.
Figure 16. 2-AA labeled glycans from fetuin analyzed with THAP/DHB as matrix in negative linear mode

Figure 17. 2-AB labeled glycans from fetuin analyzed with DHB in negative linear mode

Figure 18. 2-AB labeled glycans from fetuin analyzed with DHB/THAP as matrix in negative linear mode
7.6 EPO and follitropin

The resulting chromatograms for follitropin are shown in Figure 19a-d. Figure 19a is the chromatogram from the gradient in Table 4, 19b from Table 5, 19c from Table 6 and 19d from Table 7. The retention times for the nine main peaks are presented in Table 11. The main peaks are well resolved from each other except four peaks (5-8 in Figure 19a) that are eluting close to each other. For these four peaks the separation factor, $\alpha$, was calculated between the peaks according to Equation 5 (t$_{R1}$, t$_{R2}$ and t$_{Ro}$ are the retention times of the two peaks that are compared together with the retention time for an unretained compound). The values of $\alpha$ can be seen in Table 12. The values for $\alpha$ are consistent with the visual appearances of the spectra. A value of 1.01 for $\alpha$ indicates that the second peak has 1 % longer retention time than the first peak. The essence of the $\alpha$-values in Table 12 is that the chromatograms in figure 19c and d shows the best peak separation.

An additional way of evaluating if two peaks are baseline separated or not is to calculate the resolution, $R_s$, with Equation 6 ($w_{1,50\%}$ and $w_{2,50\%}$ is the peak width at half the peak height). These values can also be seen in Table 12. This equation takes both peak width and retention time into account, while $\alpha$ only takes the retention times into account. The resolution could not be calculated for all the peaks in Figure 19d because the peaks were not baseline separated. Since a value of 1.5 for $R_s$ is desirable, the peaks 5 and 6 together with peaks 6 and 7 are well resolved in chromatograms 19b and c (Table 12). The resolution between peak 7 and 8 is below 1.5 for both chromatograms, but the chromatogram in 19c has best separation of the peaks (Table 12). The resolution is the best possible in Figure 19c and thus these conditions is the preferable.

<table>
<thead>
<tr>
<th>Figure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>19a</td>
<td>17.8</td>
<td>18.9</td>
<td>20.6</td>
<td>21.6</td>
<td>25.9</td>
<td>26.2</td>
<td>26.6</td>
<td>26.9</td>
<td>30.7</td>
</tr>
<tr>
<td>19b</td>
<td>20.3</td>
<td>21.8</td>
<td>23.9</td>
<td>25.1</td>
<td>30.5</td>
<td>30.9</td>
<td>31.4</td>
<td>31.7</td>
<td>36.6</td>
</tr>
<tr>
<td>19c</td>
<td>33.4</td>
<td>36.8</td>
<td>42.2</td>
<td>45.1</td>
<td>58.8</td>
<td>59.6</td>
<td>61.1</td>
<td>61.7</td>
<td>74.2</td>
</tr>
<tr>
<td>19d</td>
<td>49.2</td>
<td>51.3</td>
<td>54.4</td>
<td>56.1</td>
<td>62.5</td>
<td>62.7</td>
<td>62.9</td>
<td>63.0</td>
<td>64.4</td>
</tr>
</tbody>
</table>

Table 11. The retention times for the peaks from glycans deglycosylated from fetuin from figure 19a-d.
Table 12. The calculated values for $\alpha$ and $R_s$ from Figures 19a-d.

<table>
<thead>
<tr>
<th></th>
<th>Figure 5 and 6</th>
<th>Figure 6 and 7</th>
<th>Figure 7 and 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>1.01</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>19a</td>
<td>1.01</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>19b</td>
<td>1.01</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>19c</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>19d</td>
<td>1.42</td>
<td>1.72</td>
<td>1.17</td>
</tr>
<tr>
<td>$R_s$</td>
<td>1.52</td>
<td>1.96</td>
<td>1.21</td>
</tr>
<tr>
<td>19a</td>
<td>1.82</td>
<td>3.13</td>
<td>1.43</td>
</tr>
<tr>
<td>19b</td>
<td>1.17</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

$$\alpha = \frac{t_{R2} - t_{R0}}{t_{R1} - t_{R0}} \quad \text{Eq. (5)}$$

$$R_s = 1.18 \frac{t_{R2} - t_{R1}}{w_{250\%} + w_{150\%}} \quad \text{Eq. (6)}$$

The chromatograms for EPO are shown in Figure 20a-c, a from Table 7, b from Table 6 and c from Table 2. In Figure 20a the peaks are not separated at all, but in the other two chromatograms the peaks are separated. Since the peaks have not been identified, five peaks, marked 1-5, were chosen to demonstrate how well the peaks are separated. Table 13 displays the values for $\alpha$ and $R_s$ calculated for these. The peaks 1-4 had the same value for $\alpha$. The resolution is the best, but not perfect, in Figure 20b. Peak 5 has co-eluated with a larger peak in Figure 20c and in Figure 20b the peak has begun to separate into two peaks. Of these three different gradients that were tested and evaluated, the gradient in Table 6 and Figure 29 b was the best. Looking at these peaks, they are all quite broad and this is probably an effect from using a HILIC-column with an UPLC-system. Since one of the separation mechanisms is the slow liquid-liquid extraction, the fast UPLC-separation does not make the extraction justice.
Table 13. The calculated values for $\alpha$ and $R_s$ from the Figure 29a-d

<table>
<thead>
<tr>
<th></th>
<th>Figure 1 and 2</th>
<th>Figure 3 and 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$20b$</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>$20c$</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>$R_s$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$20b$</td>
<td>1.5</td>
<td>0.81</td>
</tr>
<tr>
<td>$20c$</td>
<td>1.13</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 19. Chromatograms from UPLC-analysis of glycans deglycosylated from follitropin with four different gradients: a) gradient in table 4 b) gradient in table 5 c) gradient in table 6 and d) gradient in table 7
Figure 20. Chromatograms of glycans released from EPO analysed with different gradients. The red circles mark the peaks that were evaluated when comparing the different gradients. a) gradient in table 7 b) gradient in table 6 and c) gradient in table 2

7.7 Method validation for fetuin

Tables 14 and 15 contain the obtained values for the retention time and peak area for two peaks for fetuin. Peak 1 has a retention time around 39 minutes (Figure 9c) and the retention time for peak 2 is roughly 42.8 minutes (Figure 9c). From these values, the intra-day precision was calculated for sample 3 for day 1 and 2 (Tables 14 and 15). The value of 0.1% for the retention times is exceptionally good and the values for the relative peaks areas of 0.7 % (peak 1, Table 14) and 2.4 % (peak 2, Table 15) prove a good intra-day precision. These values prove that the UPLC-system has good precision. Table 16 contains the inter-day
precision calculated for all three samples. For the retention times the RSD is 3.2 % and 0.2 %. For the peak area the RSD is 14.0% and 34.2 %.

For the significance tests, the standard deviation of the samples was evaluated with f-tests in order to see if the standard deviations were similar. If they were similar the standard deviations could be pooled.

A significant test (t-test for comparison of two series of measurement) was done for the different injection days for sample 1 and showed that there was no significant difference between the days of injections concerning retention times, but for peak 1 the relative peak area is significantly different but not for peak 2 (Table 17).

A one-way ANOVA was performed (data not shown) between samples 1, 2 and 3, and it was seen that there was a significant difference between these. But this test does not show where the significant difference is, only that there is one. In order to evaluate where this difference was, sample 1 was compared with sample 2 and 3 in two t-tests (comparison of two series of measurements). The calculated t-values show that there is no significant difference between the retention times between sample 1 and 2, but a significant difference for the relative peak areas (Table 18). The calculated t-values for the comparison of samples 1 and 3 showed that there is a significant difference in both retention time and relative peak areas (Table 19).

All these statistical tests provide one conclusion; the method is validated for identification of glycans with UPLC but not for a quantitative measurement. It should be mention that the retention times for samples 1 and 3 is significantly different from each other and is an indication of that a change in temperature and incubation time affects the retention times. But as the peaks are well separated, there is no need to worry about that difference and the method can be concluded as robust.
Table 14. The chromatographic data from the analysis of fetuin for peak 1. For samples 1 and 2 injection 1-3 was performed on day 1. Injection 4-6 was performed on day 2. For sample 3 and 4 all of the injections were done during the same day.
*Due a calculation error there was not enough sample left in the vial before injection.

<table>
<thead>
<tr>
<th>Peak 1</th>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Injection</td>
<td>t_R(min)</td>
<td>A(%)</td>
<td>t_R(min)</td>
</tr>
<tr>
<td>1</td>
<td>39.982</td>
<td>48.1</td>
<td>39.969</td>
</tr>
<tr>
<td>2</td>
<td>39.985</td>
<td>48.7</td>
<td>39.955</td>
</tr>
<tr>
<td>3</td>
<td>39.983</td>
<td>48.3</td>
<td>39.880</td>
</tr>
<tr>
<td>4</td>
<td>39.891</td>
<td>47.9</td>
<td>39.887</td>
</tr>
<tr>
<td>5</td>
<td>39.994</td>
<td>46.2</td>
<td>39.941</td>
</tr>
<tr>
<td>6</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>\bar{x}</td>
<td>39.967</td>
<td>47.9</td>
<td>39.926</td>
</tr>
<tr>
<td>s</td>
<td>0.043</td>
<td>0.97</td>
<td>0.040</td>
</tr>
<tr>
<td>RSD(%)</td>
<td>0.1</td>
<td>2.0</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 15. The chromatographic data from the analysis of fetuin for peak 2.
For sample 1 and 2 injection 1-3 was performed on day 1. Injection 4-6 was performed on day 2. For sample 3 and 4 all of the injections were done during the same day.
*Due a calculation error there was not enough sample left in the vial before injection.

<table>
<thead>
<tr>
<th>Peak 2</th>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Injection</td>
<td>t_R(min)</td>
<td>A(µV*sec)</td>
<td>t_R(min)</td>
</tr>
<tr>
<td>1</td>
<td>42.845</td>
<td>51.9</td>
<td>42.89</td>
</tr>
<tr>
<td>2</td>
<td>42.85</td>
<td>51.3</td>
<td>42.877</td>
</tr>
<tr>
<td>3</td>
<td>42.847</td>
<td>51.7</td>
<td>42.805</td>
</tr>
<tr>
<td>4</td>
<td>42.747</td>
<td>52.1</td>
<td>42.807</td>
</tr>
<tr>
<td>5</td>
<td>42.862</td>
<td>53.8</td>
<td>42.863</td>
</tr>
<tr>
<td>6</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>\bar{x}</td>
<td>42.830</td>
<td>52.1</td>
<td>42.848</td>
</tr>
<tr>
<td>s</td>
<td>0.047</td>
<td>0.97</td>
<td>0.040</td>
</tr>
<tr>
<td>RSD(%)</td>
<td>0.1</td>
<td>1.9</td>
<td>0.1</td>
</tr>
</tbody>
</table>
**Table 16.** The calculated means, standard deviations and relative standard deviations for all of the inter-day precision for sample 1, 2 and 3.

<table>
<thead>
<tr>
<th>Inter-day precision</th>
<th>Peak 1</th>
<th>Peak 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>t_R (min)</td>
<td>A (%)</td>
<td>t_R (min)</td>
</tr>
<tr>
<td>$\bar{x}$</td>
<td>39.9</td>
<td>42.8</td>
</tr>
<tr>
<td>s</td>
<td>1.3</td>
<td>0.1</td>
</tr>
<tr>
<td>RSD (%)</td>
<td>3.2</td>
<td>14.7</td>
</tr>
</tbody>
</table>

**Table 17.** The t-values calculated for the injections from the same sample but on different days.

<table>
<thead>
<tr>
<th>Between injections</th>
<th>Peak 1</th>
<th>Peak 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>t_R</td>
<td>A</td>
<td>t_R</td>
</tr>
<tr>
<td>$t_{calc}$</td>
<td>1.54</td>
<td>1.00</td>
</tr>
<tr>
<td>$t_{crit}$</td>
<td>12.71</td>
<td>3.18</td>
</tr>
<tr>
<td></td>
<td>(p=0.05, f=1)</td>
<td>(p=0.05, f=3)</td>
</tr>
</tbody>
</table>

**Table 18.** The calculated t-values for the comparison between sample 1 and 2.

<table>
<thead>
<tr>
<th>Between sample 1 and 2</th>
<th>Peak 1</th>
<th>Peak 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>t_R</td>
<td>A</td>
<td>t_R</td>
</tr>
<tr>
<td>$t_{calc}$</td>
<td>1.54</td>
<td>0.66</td>
</tr>
<tr>
<td>$t_{crit}$</td>
<td>2.31</td>
<td>2.31</td>
</tr>
<tr>
<td></td>
<td>(p=0.05, f=8)</td>
<td>(p=0.05, f=8)</td>
</tr>
</tbody>
</table>
Table 19. The calculated values of $t$ for the $t$-test to test the robustness of the method between sample 1 and 3.

<table>
<thead>
<tr>
<th>$t_{calc}$</th>
<th>Peak 1</th>
<th>Peak 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{R}$ 6.12</td>
<td>A 56.5</td>
<td>$t_{R}$ 5.78</td>
</tr>
<tr>
<td>$t_{crit}$ 2.78</td>
<td>2.31</td>
<td>$t_{crit}$ 2.78</td>
</tr>
<tr>
<td>(p=0.05, f=4)</td>
<td>(p=0.05, f=8)</td>
<td>(p=0.05, f=4)</td>
</tr>
</tbody>
</table>

8. Conclusion
Several different experiments were performed in order to find the optimal conditions for deglycosylation of glycoproteins and the labelling and chromatographic separation of glycans. The optimized method deglycosylated the glycoprotein with a deglycosylation kit, precipitated the protein with ethanol and 60% methanol, labelled them with the non-toxic 2-AA, and used UPLC for analysis with a fluorescence detector. The 2-AA labelled showed to be more sensitive with the matrices THAP, CHCA/DHB and THAP/DHH in both negative and positive linear mode for MALDI/TOF-MS analysis compared with the 2-AB label. The validation of the method showed that the intra-day precision was excellent for both identifications and quantitative measurement, but the other significance tests proved that the method should not be used for quantitative measurements but for identification. It still remains to analyze the glycans with MALDI/TOF-MS after separation with UPLC:

9. Future aspects
As the method cannot be used for quantitative measurements, the focus must be in solving this problem in the future. A suggestion is to investigate the available kits on the market for deglycosylation, labelling and cleaning. These are easy to use and would speed up the sample preparation. Another way of improving the quantification is to use an internal standard, e.g., a stable isotopically labelled analyte. After validation, the focus should be on the collection of peaks for the MALDI/TOF-MS analysis, a possibility is to investigate automatic collection of fractions from the UPLC on already prepared targets. The stability of the labelled glycans should also be investigated; which temperature can the labelled glycans be stored at and for how long?
Changing the type and the concentration of the buffer for the separation of the glycans released from other glycoproteins should also be investigated. The temperature and length of the column could also be changed in order to influence the separation. In summary, a method for quality control of recombinant proteins is important and analyses of glycans are an excellent way of measure the quality between batches and manufacturers.

**10. Popular summary**

Monosaccharides linked together in complex patterns coupled to certain amino acids in the proteins are known as glycans. These glycans have an essential role in important functions, e.g., folding and protein activity. The patterns are influenced of the production conditions, a change in pH, nutrition or hormones will yield different patterns. These patterns can be seen as finger prints and can be used as a quality control of biopharmaceuticals like recombinant proteins.

The objectives of this thesis was to optimize parameters in a method for identifying and quantifying glycans of recombinant proteins with the help of ultra performance liquid chromatography (UPLC) with fluorescence and characterization with matrix associated laser desorption/ionisation coupled with a time-of-flight mass spectrometer (MALDI/TOF-MS). Liquid chromatography (LC) is a technique for separating analytes due to their distribution in the stationary and mobile phase. UPLC has smaller particles in the column resulting in the possibility to use higher flow rate, thus faster analysis, to get the same or even better separation compared to conventional HPLC.

MALDI/TOF-MS is a technique for analyzing the mass-to-charge ratio (m/z) of the glycans. As the charge often is one, this number equals the mass of the glycan, which thus can be identified. This sounds simple but as there are many possible combinations of monosaccharides, it is time consuming and hard to find the right glycan.

During the optimization of the method, various experiments were carried out. It was shown that glycans were release in the best way from denatured proteins together with a deglycosylation kit. As glycans do not fluoresce, the glycans had to be labelled with a molecule that works as a tag. Two different labels were evaluated with UPLC and MALDI with different matrices. The anthranilic acid label proved to be more sensitive for analysis with MALDI and the glycans had a slightly longer retention time in UPLC than glycans labelled with anthranilamide. The proteins were precipitated with two different solvents (ethanol and acetone) in order to establish which solvent was preferred. It was shown that both ethanol and acetone worked but ethanol was preferred due to a better yield.
Glycans from the glycoprotein follitropin and the recombinant protein EPO from the biopharmaceutical EPREX® were analysed with different separation conditions for the UPLC in order to find the best conditions for separation. Acceptable separation conditions for glycans released from follitropin were found, however, for EPO the case was more difficult and optimal separation conditions could not be found. In the end, a validation was done for the method for fetuin. The inter- and intra-day precision together with the robustness were evaluated and showed that the method could be used for identifications but not for quantitative measurements.
References

Appendix A

Mass spectra from glycans labelled with 2-AB

**DHB**

Positive mode

**CHCA**

Negative mode

Positive mode
**THAP**

Negative mode

Positive mode

**THAP/CHCA**

Negative mode

Positive mode
**DHB/CHCA**

Negative mode

Positive mode

---

**DHB/THAP**

Positive mode
Mass spectra from glycans labelled with 2-AA

**DHB**

Negative mode

Positive mode

**CHCA**

Negative mode

Positive mode
**THAP**

Positive mode

**THAP/CHCA**

Negative mode

Positive mode
**DHB/CHCA**

Positive mode

![DHB/CHCA spectrum](image)

**DHB/THAP**

Positive mode

![DHB/THAP spectrum](image)