SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Towards Next Generation Glycoanalytics: Automatable Biomarker Discovery

by Csaba Váradi

Supervisor: Prof. András Guttman

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF MOLECULAR MEDICINE

DEBRECEN, 2015

Towards Next Generation Glycoanalytics: Automatable Biomarker Discovery

By Csaba Váradi, Molecular Biology (MSc)

Supervisor: Prof. András Guttman, PhD, DSc

Doctoral School of Molecular Medicine, University of Debrecen

Head of the **Examination Committee**:	Prof. János Szöllősi, PhD, DSc
Members of the Examination Committee:	Dr. János Kerékgyártó, PhD
	Dr. Szabolcs Sipeki, PhD

The Examination takes place at Department of Biophysics and Cell Biology, University of Debrecen, December 16, 2015, at 11 a.m.

Head of the **Defense Committee**:	Prof. János Szöllősi, PhD, DSc
Reviewers:	Prof. Éva Szökő, PhD, DSc
	Dr. Attila Gáspár, PhD
Members of the Defense Committee:	Dr. János Kerékgyártó, PhD
	Dr. Szabolcs Sipeki, PhD

The PhD Defense takes place at the Department of Obstetrics and Gynecology, Faculty of Medicine, of University of Debrecen, December 16, 2015, at 1 p.m.

1. Introduction

The well-known aspect of carbohydrates is their central role in energy metabolism, although in the last few years increased attention was given for oligosaccharides as post-translational modifications. The process when one or more carbohydrate chains are added to proteins which are assembled in the endoplasmic reticulum called glycosylation, the attached sugars are glycans. Glycosylation is a principal chemical modification of proteins, responsible for physical, chemical and effector functions of proteins. The attached glycans are composed of monosaccharide units which can be combined in many different ways due to the linkage type and position. The main aspect of glycobiology is to provide structural information of carbohydrates attached to proteins and reveal how these oligosaccharides can mediate biological processes. The function of the attached glycans can be very diverse usually separated into two groups:

1. Intrinsic functions performed by glycans:
    a. Structural components of cell walls and extracellular matrix
    b. Modifying solubility and stability of proteins
2. Extrinsic functions by lectin-glycan interactions:
    a. Mediating and modulating cell-cell and cell-matrix interactions
    b. Intra- and extracellular trafficking of glycoconjugates
    c. Intra- and extracellular signaling

There are two main glycosylation types based on the linkage of the glycan to the parent protein. N-linked glycans are attached through a nitrogen atom to an asparagine while O-linked glycans are connected through oxygen to a serine or threonine. O-linked glycans are usually smaller structures, consisting of a few monosaccharide units while N-links are more complex containing several branches, in some cases terminating with a negatively charged sialic acid residue. Both types of glycosylation can be found on membrane proteins and also on glycoproteins which are secreted in biological fluids. The lack of specific enzymes for most O-type glycan release resulted that N-linked glycosylation is more widely studied although their biological importance is considered to be equal. Both glycosylation types are essential in proper folding, stability and the effector functions of the parent protein. The primary polypeptide sequence of proteins is directly encoded in the genes; however, the glycan structures are built up by monosaccharide-

and linkage-specific glycosyl-transferases, thus glycans are considered as secondary gene products. Although the origin of glycans is non-template driven, the structures of the specific enzymes, which are responsible for construction of the attached carbohydrates, are coded in the genome. This was first proven by knock-out mice based techniques showing that the elimination of glycoconjugates can lead to health disorders. The carbohydrate content of glycoconjugates is variable, as the number of glycosylation sites and the occupancy at particular sites can vary within a protein and also from proteins to proteins. For example in antibodies it is relatively low (<5% of total mass), while other glycoproteins can be highly glycosylated such as haptoglobin (19%), erythropoietin (40%), α1-acid glycoprotein (45%), mucins (40-80%). Glycoproteins can have multiple glycosylation sites with different glycoforms on each site resulting broad site-specific structural diversity referred to as microheterogeneity of protein glycosylation. The vast structural heterogeneity of glycans is based on the expression level and activity of glycosidases and glycosyl-transferases, which can be altered in pathological conditions therefore, comprehensive glycosylation analysis offers a new avenue in biomarker discovery as glycosylation changes can be sensitive indicators of the actual state of the underlying biochemical mechanism. With the help of efficient glyco-markers, diseases may be diagnosed at early stages, also saving patients from sometimes complicated and invasive diagnostic procedures. The recognition of clinical relevance of glycosylation changes resulted the recent golden age of Glycomics.

2. Aims and Motivation

Glycosylation is an essential post-translational modification of proteins with effects on their physical, chemical properties and effector functions. The recognition of clinical relevance of glycosylation changes in pathological conditions resulted in increased attention towards glycomics, although the advent of monoclonal-antibody therapeutics has given an even greater impetus to this field of research. Glycosylation analysis was applied in clinical diagnostics as disease marker discovery of alcohol abuse and hepatocellular carcinoma proving the biomarker potential of this protein modification.

Altered haptoglobin glycosylation was reported as candidate marker in different lung diseases. Our aim was full structural annotation of haptoglobin glycans and their quantitation in normal, acute inflammatory, chronic inflammatory and malignant conditions focusing on branching- and fucosylation-degree.

Another widely studied potential biomarker is the glycosylation of immunoglobulin G. Glycosylation variations of human serum IgG have been found in conditions such as ageing, pregnancy, inflammatory diseases and cancer. One of the most common alterations in IgG glycosylation is reduced terminal galactosylation, and the concomitantly amount of GlcNAc terminating structures in inflammatory diseases. One of our goals was to examine if IgG glycosylation was a useful marker for monitoring treatment efficacy and identifying responders. The reportedly decreased galactosylation of IgG in RA and CD was expected to be altered towards to the normal profile after treatment in correlation with the responder status of the patients.

The increasing use of glyco-markers and the huge market of therapeutic monoclonal antibodies necessitate high-throughput and automatable methods for precise glycan annotation. Currently used methods require days for sample preparation using expensive laboratory equipment and the use of centrifugation technique that is hard to automate. A new method was aimed to be developed and implemented to alleviate the need of vacuum-centrifugation and overnight incubations, thus easily automatable and needs only hours instead of days for sample preparation. This would give the opportunity of automated sample preparation for biomedical and biopharmaceutical research and development.

3. Materials and Methods

3.1. Chemicals and Reagents

Water and acetonitrile were Chromasolv HPLC grade from Sigma Aldrich (St. Louis, MO, USA). Haptoglobin standard, dithiothreitol, iodoacetamide, acetic acid, and all other chemicals were also purchased from Sigma Aldrich. High purity 8-aminopyrene-1,3,6-trisulfonate (APTS), Agencourt Cleanseq magnetic beads were from SCIEX (Brea, CA, USA). Purified N-linked carbohydrate standards, including bi-, tri-, tetra-antennary glycans (A2, A3, A4), their analogues carrying terminal galactose residues (A2G2, A3G3, A4G4) and the core fucosylated bi-antennary galactosylated glycan (FA2G2), CU clean-up cartridges were obtained from Prozyme (Hayward, CA, USA). 96 Protein A deep-well plate was purchased from Thermo–Fisher (Waltham, MA, USA). The PNGase F deglycosylation kit was purchased from New England Biolabs (Ipswich, MA, USA) and the digestion reaction mixtures were prepared following the manufacturer's protocol.

3.2. Haptoglobin Purification from Plasma Samples

Protein-G chromatography (GE-Healthcare, Bucks, UK) was used for purification of the haptoglobin specific mAbs from Biosystems International (Debrecen, Hungary). 10 mg of haptoglobin specific mAb were coupled to 0.5 g CNBr-activated Sepharose 4B affinity beads according to the manufacturer's protocol (GE-Healthcare). Plasma from age and gender matched human control (2 males, average age 61.0), pneumonia (3 males, average age 60.3), COPD (3 males, average age 61.6) and lung cancer (3 males, average age 61.3) groups were collected in the Department of Pulmonology at the University of Debrecen (Hungary) under clinical protocol RKEB /IKEB:2422-2005. The study was approved by the Hungarian Research Ethics Committee and all participants signed informed consent forms. 200 µL of individual patient plasma samples were diluted 5 times with 50 mM potassium hydrogen phosphate (pH 7.0) binding buffer. HSA was depleted by incubation with 250 µL Blue Sepharose 6 fast flow slurry (GE-Healthcare) for 30 min at room temperature. Following centrifugation, the albumin depleted flow-through was incubated with 250 µL Protein-G4 fast flow slurry (GE-Healthcare) for 30 min at room temperature to deplete the IgG content. HSA and IgG depleted plasma was loaded onto an in-house made mAb affinity column (1 ml bed-volume) and incubated for 1 hour at room temperature. The column was washed with 7 column volumes of phosphate-buffered saline (pH

7.2) and haptoglobin was eluted with 2 x 1 ml 0.2 M glycine - HCl elution buffer (pH 2.7). Eluted haptoglobin was immediately neutralized by the addition of 0.2 ml of 1 M Tris-HCl (pH 9.0) buffer. Protein concentration was measured by bicinchoninic acid (BCA) kit according to the manufacturer's protocol (Pierce, Rockford, IL, USA). SDS–PAGE on 4-20% Tris-Glycine precast gradient gels (Invitrogen, Carlsbad, CA, USA) were used to confirm the purity of haptoglobin.

3.3. Glycan Release and Derivatization of Haptoglobin Samples

Haptoglobin samples (approximately 50 μg/sample) were reduced with dithiothreitol, alkylated with iodoacetamide and dissolved in 50 μL of 50 mM sodium-bicarbonate (pH 7.0). Release of the N-linked glycans was accomplished by the addition of 2 U of recombinant PNGase F (Prozyme) and incubation at 37°C overnight. The released glycans were separated from the remaining polypeptide chains by means of 10 kDa cut-off spin filters (Nanosep 10 kDa, Sigma-Aldrich) and dried in a centrifugal vacuum evaporator (Thermo Scientific, Asheville, USA). The dried sugars were fluorescently labeled via reductive amination by the addition of 5.5 μL of 20 mM APTS in 15% v/v acetic acid and 1.5 μL of 1 M sodium cyanoborohydride in tetrahydrofuran, overnight at 37°C. The reaction was stopped by the addition of 93 μL of water and the unreacted fluorescent dye was subsequently removed by normal phase resin containing pipette tips (PhyNexus, San Jose, CA, USA).

3.4. Patient Samples for IgG glycosylation and Transcriptome Analysis

The Institutional Review Board of University of Debrecen Medical and Health Science Center approved the clinical protocol and the study in compliance with the Helsinki Declaration. Signed informed consent was obtained from all individuals providing blood sample. RA patients were on maximal-tolerable methotrexate treatment (5 to 30 mg per week), which was stable for at least 4 weeks before baseline. Prednisone therapy (≤10 mg per day) was allowed for CD patients, which had to be stable for at least 2 months before infliximab therapy.

3.5. Sample Preparation and Fluorophore Labeling of IgG Glycans

50 μL of serum samples were diluted with 250 μL PBS buffer (pH 7). The diluted samples were loaded into the wells of a Protein A microwell plate (Thermo Scientific), washed with 300 μL PBS and the captured proteins were eluted by 200 μL of 10% acetic acid. 10 kDa Nanosep

Spinfilters (Sigma) were used for buffer exchange. N-glycan release was performed on the 10 kDa spinfilters after the proteins were reconstituted by 50 µL of 50 mM ammonium-bicarbonate. For PNGase F digestion the Prozyme deglycosylation kit was used according to the manufacturer's protocol. The released glycans were dried in a vacuum-centrifuge and then fluorescently derivatized by the addition of 6 µL of 20 mM APTS in 15 % acetic acid and 2 µL of 1 M sodium-cyanoborohydride using overnight incubation at 37 °C. The labeled samples were purified using Prozyme CU cartridges.

3.6. Capillary Electrophoresis

Capillary electrophoresis profiling of the APTS labeled N-glycans was performed in a PA800 *Plus* automated capillary electrophoresis instrument (SCIEX) equipped with a solid state laser based fluorescent detector (excitation: 488 nm, emission: 520 nm). The separations were accomplished in 50 cm effective capillary length (60 cm total) N-CHO coated 50 µm i.d. capillary columns (SCIEX) filled with the N-CHO Carbohydrate Separation Buffer (SCIEX). The applied electric field strength was 500 V/cm, with the cathode at the injection side and the anode at the detection side (reversed polarity). Samples were injected by pressure: 1 psi (6.89 kPa) for 5 sec. For migration time correction and relative quantification purposes, APTS labeled maltose (G2) was co-injected with each sample as internal standard. The Karat 32 version 7.0 software package (SCIEX) was used for data acquisition and analysis.

3.7. Gene Expression Analysis

Gene expression data was derived from global transcriptome analysis of the patient cohorts. Microarray data are available in Gene Expression Omnibus database (accession number: GSE42296).

3.8. Statistical Analysis

IBM SPSS 20 software (New York, US) was used to perform Mann-Whitney pairwise comparison for all statistics. P values of <0.05 were considered to be significant.

4. Results

4.1. Increased Branching- and Fucosylation-Degree of Haptoglobin N-Glycosylation in Inflammatory and Malignant Lung Diseases

After albumin and IgG depletion the haptoglobin was partitioned from the plasma samples of patients by affinity purification. The purity of the partitioned haptoglobin was checked by SDS-PAGE. N-glycans were then released by PNGase F, fluorescently labeled with APTS and analyzed by CE-LIF after the removal of the excess dye. Using the relative percentage area values of the desialylated glycan structures of haptoglobin, fucosylation and branching degrees were determined in the control, acute and chronic inflammation as well as cancer patient samples. Core (FDc) and arm (FDa) fucosylation degrees were evaluated separately. Alterations in size of the glycans were examined based on branching degree (BD) i.e., upon the level of various bi- (A2G2, A2FG2, FA2G2), tri- (A3G3, A3FG3, FA3G3) and tetra- (A4G4, A4FG4, FA4G4) antennary structures.

4.1.1. Branching degree

Altered branching of the glycosylation of serum glycoproteins has been observed in pathologic conditions, thus, branching degree (BD) was investigated to identify the changes in the antennary structures between the different groups.

Branching Degree (BD) = (Tri+tetra-antennary glycans) / (Total glycans) Eq 1.

Elevated branching was found in all three disease groups with COPD showing the highest value compared to the control. A significant decrease was found in the amount of A2G2 glycan and an increase in A3G3 glycan contributing to the elevated branching degree of this group.

4.1.2. Fucosylation degree

The two types of fucosylation on human haptoglobin are core-fucosylation, where the fucose unit is linked (α1-6) to the reducing end GlcNAc residue that is directly connected to the polypeptide backbone, and antennary-fucosylation where the fucose unit is linked to any of the antennary GlcNAc residues via α1-3 and/or 4 linkages. The degree of fucosylation (FD) can be expressed as the amount of fucosylated glycans relative to the total amount of glycans.

Fucosylation degree (FD) = (Fucosylated glycans) / (Total glycans) Eq 2.

Total fucosylation degree was calculated for the control, pneumonia, COPD and lung cancer patient groups using equation 2. The different fucosylation type containing structures (i.e., antennary and core) were also analyzed separately and referred to as core (FDc) and arm fucosylation (FDa) degrees. In total fucosylation, a slight decrease was observed in pneumonia and COPD and a strong increase in lung cancer. The core and arm fucosylation ratio revealed that the decrease in total fucosylation degree in pneumonia and COPD was the results of the interplay between the significant decrease in antennary-fucosylation and the slight increase in core-fucosylation, i.e., the elevated core fucosylation degree could not compensate for the arm fucosylation change. This observation suggested that it is critical to measure both the core and arm fucosylation degrees individually because it is possible that while the total fucosylation degree is apparently unchanged, the subgroups of FDc and FDa are potentially altered due to different expression of the fucosyl-transferases transporting the α1-6 core and α1-3/4 arm fucose units. Please note that the highest degree of total fucosylation was observed in the lung cancer group, however a closer look revealed that the arm-fucosylation was highly elevated in lung cancer, suggesting that the increase of the α1-3/4 linked arm fucosylation was the reason of the higher total fucose content in that group.

4.2. Combination of IgG N-Glycosylation and Corresponding Transcriptomics Data to Identify Anti-TNFα Treatment Responders in Inflammatory Diseases

The aim of this study was to investigate the changes of serum IgG glycoforms in Crohn's disease and Rheumatoid arthritis patients in response to anti-tumor necrosis factor alpha (TNFα) treatment. IgG was isolated from patient serum samples using Protein A affinity pull-down, followed by the release of N-glycans with PNGase F. The released glycans were fluorescently tagged with aminopyrene-trisulfonate and analyzed by capillary gel electrophoresis with laser induced fluorescent detection. Utilizing the high resolution of CE-LIF, 26 IgG glycoforms were examined in order to reveal significant changes in response to anti-TNFα treatment.

4.2.1. IgG N-glycosylation analysis before and after anti-TNFα treatment in RA

Special assessment was developed for RA to follow disease remission and determine patient responsiveness for the therapy. In RA, DAS28 (disease activity score examining 28 joints, which are commonly affected by RA) is the measure of disease activity in combination with measuring the level of C reactive protein (CRP). Due these long and complicated diagnostic procedures,

there is a need to find non-invasive ways to determine markers that readily differentiate non-responders and responders in an early stage of the treatment. 17 serum samples from RA patients were analyzed. Six (5 females, 1 male, average age 44.3 years) were identified as responders and 11 (9 females, 2 males, average age 47 years) as non-responders by clinical evaluation based on the DAS28 scoring system. Three low (>2%) abundant structures of A2(3)G1, A2B(3)G1 and FA2B(3)G1 were found to be significantly different between responders (R) and non-responders (NR) of RA patients before anti-TNFα treatment. All of the significantly altered structures were galactosylated and in each case, responders showed higher galactosylation level. However, there were no significant changes in response to anti-TNFα therapy in any of the RA subgroups. It has to be noted that in both subgroups (R, NR) higher galactosylation (consequently lower agalactosylation) was detected after the therapy but this change was minimal. These findings are in agreement with previous studies where IgG galactosylation was reportedly altered in RA patients but failed to predict clinical response after anti-TNFα treatment. To understand the molecular background of the observed glycosylation differences, the expression levels of the corresponding glycosyl-transferases and glycosidases were also investigated. In agreement with the glycosylation analysis data, there were no significant changes in the expression level of the respective enzymes after the treatment, although one reasonable difference was found between responders and non-responders at baseline (before treatment). Higher galactosidase activity was found in non-responders suggesting the reason of the detected lower galactosylation level compared to responders.

4.2.2. Differences of IgG glycosylation before and after anti-TNFα treatment in CD

In CD, the combination of different clinical diagnostic indicators are used to quantify the symptoms, such as number of liquid stools, abdominal pain, general well-being, extraintestinal complications, need of antidiarrheal drugs, abdominal mass, hematocrit and body weight all summarized in Crohn's Disease Activity Index (CDAI). Based on the CDAI, 14 patients were identified as responders (R, 6 females, 8 males, average age 36.2 years) and 5 as non-responders (NR, 3 females, 2 males, average age 36 years). The serum samples were taken before and 2 weeks after the treatment, although the identification of responders was only possible after months by the CDAI protocol. Early recognition of status response could prevent patients from ineffective medication so our aim was to find one of the high abundant IgG glycoforms possibly providing useful information in responsiveness.

Similarly to RA, even before the start of the treatment, significant differences were revealed between the two subgroups (R vs NR). Using Mann-Whitney pairwise comparison, A2G2S1, FA2BG2S1, FA2(3)G1 and FA2BG2 were found significantly different although it has to be noted that among these, only the FA2(3)G1 was not a low abundant structure. It has been reported earlier that serum IgG galactosylation level correlated with the disease activity in CD. Our results are in agreement with these findings as non-responders showed significantly lower galactosylation (FA2(3)G1) level suggesting altered activity of galactosyl-transferases/galactosidases in responders and non-responders. However, we have not found significant differences in the expression level of the corresponding enzymes before the treatment.

The main goal of this study was to find glycosylation markers that were significantly altered in response to anti-TNFα treatment and thereby identify responders. We have successfully found an IgG glycoform, which level was significantly altered 2 weeks after the therapy. In the responder group of CD, FA2G2S1 was found to be significantly ($p=0.01$) higher 2 weeks after anti-TNFα treatment. Sialylation was found to be important earlier in the anti-inflammatory activity of immunoglobulins, which support our findings as the higher anti-inflammatory activity of IgGs can be helpful in inflammatory disease. To confirm our findings, the expression levels were determined for the relevant sialyl-transferases and sialidases. Significant differences were found only in the responder group of CD where higher sialyl-transferase activity and lower sialidase activity was detected 2 weeks after the infliximab treatment. These results are in partial agreement with previous reports, where higher sialidase activity was found in CD and ulcerative colitis. The transcriptomics data readily supported our IgG glycosylation based findings as the combination of higher sialyl-transferase and lower sialidase activity might be responsible for the higher sialylation level of the generated glycoproteins.

4.3. Magnetic Bead Based Sample Preparation for Automated N-glycan analysis

In this part of our work, an optimized, rapid and high-throughput magnetic bead based sample preparation approach was developed for N-glycosylation analysis, where all preparation steps can be easily automated using simple liquid handling robots. Particular attention was given to avoid any centrifugation steps (both normal and vacuum) and overnight incubations, which are otherwise part of regular protocols of most glycan sample preparation methods. Utilization of the magnetic bead based approach alleviated the need of any centrifugation or centrifugal vacuum evaporation steps. The sample preparation protocol was demonstrated via practical examples using representative glycoprotein standards possessing high mannose, complex and highly sialylated glycans. All individual processing steps, such as glycan release, fluorescent labeling and fluorophore label cleanup were optimized to decrease processing time and increase the efficiency for the magnetic bead based method.

4.3.1. Glycan release

Minimizing PNGase F mediated deglycosylation time was important, while considering easy implementation of this part for a magnetic bead based sample preparation platform. Using the liquid handling robot friendly open 96 well-plate format, the effect of temperature was important to consider as >60°C could cause protein denaturation and buffer evaporation, this latter especially in low volumes (10-50 µL).

Digestion efficiency was compared at 50°C and 37°C for the deglycosylation of IgG, fetuin and RNase B glycoprotein standards using 0.5, 1, 2, 4, 8 and 16 hours of incubation with the goal to find the shortest time that results in full deglycosylation. The released glycans were APTS labeled and analyzed by CE-LIF. Three releases were made with each digestion strategies and three analyses were made with each release, generating nine data points per digestion time and temperature. While peak distribution (measured as peak area percentages) exhibited no differences between the two incubation temperatures, the RFU values represented the changes in the amount of the released glycans. Peak intensities increased significantly slower at 37°C comparing to 50°C, where the maximum level was already reached after 1 hour of incubation time. In addition, no differences were found in peak area percentages comparing the various digestion times, including the overnight reaction, suggesting that the same glycosylation pattern was obtained using shorter incubations with the main difference in the amount of the released

sugars. As expected, accelerated glycan release was observed at higher temperature, thus 1 hour/50°C PNGase F digestion was used in all consequent experiments.

4.3.2. APTS labeling

The goal of the labeling optimization part was to identify conditions that would support magnetic bead based automation using simple liquid handling robots, while providing the same labeling efficiency as conventional centrifugation based methods, in special respect to peak intensity and peak area distribution. Another important aspect of this part of the work was to achieve high labeling efficiency without the need of overnight incubation and vacuum-centrifugation based sample concentration. First, mono- and di-sialo glycan standards of A2G2S1, FA2G2S1, A2G2S2 and FA2G2S2 were labeled in duplicates with 20 mM APTS in 15 % acetic acid for 2 hours at 37, 50, 65 and 80°C. The non-sialylated counterparts of these glycans (A2G2 and FA2G2) were also labeled and used for spiking the higher temperature reaction mixtures to reveal possible temperature induced desialylation. The increase in the reaction temperature significantly elevated the desialylation process for all sialylated glycan standards. Interesting to note that di-sialo standards exhibited more enhanced sialic acid loss. On average, 2% sialic acid loss was observed at 50°C, 11% at 65°C and 33% at 80°C, suggesting that carefully chosen derivatization temperature is crucial during glycan labeling when sialylated structures are expected in the sample. The effect of incubation time on fluorophore labeling efficiency was examined at 37°C. Significant differences were observed between overnight and 2 hour long incubation. The mono- and di-sialo structures showed on average 2-3% and 9-10% sialic acid loss, respectively, using overnight incubation, signifying the importance of incubation time length in sialylation loss. While shorter APTS labeling times (2 h) resulted in lower signal intensity, longer (overnight) incubation caused noticeable sialic acid loss. Considering the fact that reliable sample preparation and data generation is more important than signal intensity, a 2 hour labeling at 37°C was chosen as a good compromise and was used in all consequent experiments.

In an effort to compensate for the lower signal intensity at shorter incubation times, the effect of catalyst (acetic acid) and fluorophore (APTS) concentration was also revisited. Using 20 mM APTS in 15, 20 and 25% acetic acid, mono- and di-sialo glycan standards (A2G2S1, FA2G2S1, A2G2S2, and FA2G2S2) were labeled to assess the labeling efficiency, while trying to minimize

sialylation loss. With the use of 20% acetic acid, the resulted peak pattern and area percentages showed no detectable sialic acid loss, while significantly higher peak intensities were observed. The effect of APTS concentration on the labeling efficiency was examined by maltooligosaccharide ladder standard using 20, 40 and 80 mM APTS in 20% acetic acid at 37°C for 2 hours (triplicates). The use of higher APTS concentration lead to increased labeling efficiency. In spite of the fact that the use of 80 mM APTS resulted in the highest intensity, a compromise was suggested here too to use 40 mM APTS in the higher volume labeling reaction that was required by automation consideration (minimum of 20 µL volume). Using the above suggested dye and catalyst concentrations, released glycans from 100 µg IgG, fetuin and RNase B glycoprotein standards were labeled in duplicates using 40 mM APTS in 20% acetic acid at 37°C for 2 hours and compared to the traditionally used 20 mM APTS in 15% acetic acid labeling strategy with 2 hour and overnight incubation. The combination of higher dye and catalyst concentrations resulted in approximately 20% higher labeling efficiency compared to the original 2 hour labeling strategy without any apparent sialic acid loss. Please note that signal intensity was still >50% lower that with the overnight reaction where on the other hand noticeable sialic acid loss was detected. As a first approximation we suggest APTS labeling at 37°C for 2 hours with 40 mM APTS in 20% acetic acid obtain adequate labeling efficiency and still minimizing sialic acid loss.

4.3.3. Magnetic bead based sample preparation

As emphasized earlier, the key aspect of this part of the work was to find optimal PNGase F based glycan release and APTS labeling parameters for easy downstream automation, with the key aspect to avoid any centrifugation steps during the sample preparation process, including vacuum centrifugation. Carboxyl coated magnetic beads were apparently capable of capturing complex carbohydrates both in their native (after glycan release) and fluorophore labeled (after APTS labeling) forms. First, magnetic beads were applied to clean-up the APTS reaction mixture, i.e., to remove excess APTS, not conjugated to the sugar structures. APTS labeled hIgG, fetuin and RNase B glycans were purified in triplicates after the labeling reaction using 200 µL carboxyl coated magnetic bead suspension. Binding and washing steps were accomplished by using 150 µL of 87.5% acetonitrile. The elution step was accomplished with the use of 25 µL of water. Please note that ≥20 µL volumes of magnetic bead suspension and binding/elution solutions were readily handled by automatic pipettors, suggesting the utilization

of the same volumes in simple liquid handling robots with regular pipette tips or syringes. The eluate was directly analyzed by CE-LIF without any further processing. Second and third elution fractions were also analyzed to assess the efficiency of the first elution. When the cleanup mixture was suspended properly and 25 μL of water was used for elution, no detectable sample remained on the beads, i.e., the second and third elution gave negative results. On the other hand, when only 15 μL of water was used in the first elution, traces of remaining APTS labeled glycans were detected. No differences were found; however, in peak area distribution using the magnetic bead based cleanup protocol, i.e. the resulting profile was identical to the one that was obtained with conventional sample cleanup methods, suggesting no apparent bias for the different glycan structures (neutral, sialylated, high mannose) towards the beads, while most of the free APTS was removed for IgG, RNase B and fetuin glycans, respectively.

Based on the successful implementation of centrifugation free APTS cleanup protocol with the use of a carboxyl coated magnetic beads, a similar approach was evaluated to capture the released glycans after PNGase F digestion. Magnetic beads were added to the PNGase F reaction mixture after the incubation step, first to bind the released glycans in 87.5% acetonitrile solution. In this case, however, instead of water, the free glycans were eluted by aqueous APTS solution (40 mM in 20% acetic acid) followed by the addition of the reducing agent (1 M pic-BH3 in MeCN) to immediately initiate the labeling reaction without any interim pre-concentration steps. Please note that this approach, while very effective, did not require any vacuum centrifugation based sample pre-concentration or any other purification steps to remove the remaining polypeptide chain and PNGase F enzyme from the digestion reaction mixture.

To conclude this section, a magnetic bead based protocol was developed for N-glycosylation analysis of glycoproteins not requiring hard-to-automate centrifugation and vacuum-centrifugation steps. Glycan release, APTS-labeling and clean-up were optimized resulting in 4 hours magnetic bead based process with excellent yield and high reproducibility. The next step of this work is to apply this optimized magnetic bead based protocol with all steps from PNGase F digestion, through optimized fluorophore labeling and clean-up to high throughput sample processing in 96 well plates format with simple liquid handling robots allowing full automation.

5. Discussion

5.1. Preliminary Findings on Haptoglobin N-glycome Alterations

In our experiments disease associated changes of core and antennary fucosylation were identified by targeted exoglycosidase digestions and their levels were compared in the different patient groups. Terms such as of core- and arm-fucosylation degree, as well as branching-degree were introduced for easier characterization of the changes. Increased level of α1-6 fucosylated tri-antennary glycans was found in all disease groups compared to the control and also elevated amounts of core- and arm fucosylation on tetra-antennary glycans were detected in the lung cancer compared to the COPD. An increased level of tri- and tetra-antennary glycans exhibiting the Sle^{x} epitope on haptoglobin in chronic inflammatory conditions, reportedly suggested a pro-inflammatory impact on the mechanism of glycosylation. Ang et al. found increased serum concentration of α1-6 fucosylated haptoglobins in hepatocellular carcinoma compared to chronic liver disease while Sarrats et al. reported increased arm-fucosylation on haptoglobin in chronic pancreatitis, while in pancreatic cancer it was just slightly elevated. It has been also described that core-fucosylation on haptoglobin was increased in stage IV pancreatic cancer. Based on these results they concluded that Sle^{x} and increased branching were probably associated with the inflammatory response, while the increased core-fucosylation on acute phase proteins might be cancer specific. These results suggest the importance of the analysis of different fucosylation types in various diseases which is in agreement with our findings.

5.2. IgG Glycosylation Changes in Response to Anti-TNFα Treatment

IgG glycosylation is reportedly a reversible alteration in inflammatory diseases. During biological therapies, decreased galactosylation was found to be altered towards to the normal healthy profiles. Ercan et al reported that serum hypogalactosylation fails to predict patient responsiveness in RA 3 months after the therapy although galactosylation level was improved with the treatment. These findings are in agreement with our results as there was no significant glycosylation alteration in response to the treatment in RA but an increased galactosylation was detected in both responder and non-responder group. Similar results have been published in CD where the agalactosyl IgG level was correlated with the disease activity, suggesting high predictability of therapeutic outcome. Our results are in partial agreement with these findings, as

lower galactosylation was detected in CD although similarly to RA it failed to predict patient response. Using intravenous immunoglobulins, sialylation was found to be crucial in anti-inflammatory activity of IgG-s, while terminal galactosylation was associated with CDC activation. These observations suggest that in inflammatory diseases, IgG sialylation might have higher biomarker potential than galactosylation which supports our results as higher IgG sialylation was found in CD after anti-TNFα therapy, confirmed by transcriptomics data analysis of the corresponding glycosyl-transferases and glycosidases

5.3. Novel Magnetic Bead Based Glycan Preparation

Full automation to enable high throughput N-glycosylation profiling and sequencing with good reproducibility is critical to fulfill contemporary needs of the biopharmaceutical industry and requirements of national regulatory agencies. The most prevalently used glycoanalytical methods are capillary electrophoresis and hydrophilic interaction liquid chromatography, while very efficient, both necessitate extensive sample preparation and cleanup, including glycoprotein capture, N-glycan release, fluorescent derivatization, purification and pre-concentration steps during the process. Currently used protocols to fulfill these tasks require multiple centrifugation and vacuum-centrifugation steps making liquid handling robot mediated automated sample preparation difficult and expensive. We reported a rapid magnetic bead based sample preparation approach that enables full automation including all the process phases just in 4 hours without requiring any centrifugation and/or vacuum concentration steps. Please note that the applied carboxylated magnetic beads were never used in glycan preparation before. This novel magnetic bead based protocol has been compared to conventional centrifugation based sample preparation strategies using standard glycoproteins (IgG, fetuin and RNase B) and featured rapid processing time, high release and labeling efficiency, good reproducibility and the potential of easy automation.

6. Summary

The biomarker potential of glycosylation is presented in this work focusing on haptoglobin and IgG. Haptoglobin glycosylation was reportedly altered in inflammatory and malignant conditions. Fucosylation- and branching-degrees were examined and the importance of linkage-specific fucosylation-degrees was pointed out. The preliminary results were in agreement with the literature as increased branching and fucosylation was found in various pathological conditions.

Prediction of patient response for any therapy is critical in inflammatory diseases such as CD and RA, as currently used scoring systems require months for responder identification. The importance of reliable biomarkers in these diseases is essential as the efficacy of biological therapies can vary between patients. Utilizing the high resolving power of CE-LIF, 26 IgG glycoforms were examined in RA and CD before and after anti-TNFα treatment. In RA, three low abundant galactosylated structures were found to be significantly different before the treatment and all responders showed higher galactosylation levels. No significant alteration was detected in RA in response to the treatment. In CD FA2G2S1 level was significantly increased in response to anti-TNFα therapy, thus being a possible candidate marker for responder identification. This finding was also supported by transcriptomics analysis of the corresponding glycosyl-transferase and glycosidase enzymes, as higher sialyl-transferase and lower sialidase activity were found.

For reliable biomarkers huge number of samples has to be analyzed which necessitates automated high-throughput methods. A magnetic bead based protocol was developed for N-glycosylation analysis of glycoproteins not requiring hard-to-automate centrifugation and vacuum-centrifugation steps. Glycan release, APTS-labeling and clean-up were optimized resulting in a 4 hours magnetic bead based process with excellent yield and high reproducibility. The next step of this work is to apply this optimized magnetic bead based protocol with all steps from PNGase F digestion, through optimized fluorophore labeling and clean-up to high throughput sample processing in 96 well plates format with simple liquid handling robots allowing full automation.

UNIVERSITY OF DEBRECEN
UNIVERSITY AND NATIONAL LIBRARY

deenk

Registry number: DEENK/98/2015.PL
Subject: Ph.D. List of Publications

Candidate: Csaba Váradi
Neptun ID: MH5HSS
Doctoral School: Doctoral School of Molecular Medicine
MTMT ID: 10037921

List of publications related to the dissertation

1. **Váradi, C.**, Holló, Z., Póliska, S., Nagy, L., Szekanecz, Z., Váncsa, A., Palatka, K., Guttman, A.: Combination of IgG N-glycomics and corresponding transcriptomics data to identify anti-TNFα treatment responders in inflammatory diseases.
   Electrophoresis. Epub ahead of print (2015)
   DOI: http://dx.doi.org/10.1002/elps.201400575
   IF:3.161 (2013)

2. **Váradi, C.**, Lew, C., Guttman, A.: Rapid Magnetic Bead Based Sample Preparation for Automated and High Throughput N-Glycan Analysis of Therapeutic Antibodies.
   Anal. Chem. 86 (12), 5682-5687, 2014.
   DOI: http://dx.doi.org/10.1021/ac501573g
   IF:5.825 (2013)

3. **Váradi, C.**, Mittermayr, S., Szekrényes, Á., Kádas, J., Takács, L., Kurucz, I., Guttman, A.: Analysis of Haptoglobin N-glycome Alterations in Inflammatory and Malignant Lung Diseases by Capillary Electrophoresis.
   Electrophoresis. 34 (16), 2287-2294, 2013.
   DOI: http://dx.doi.org/10.1002/elps.201300041
   IF:3.161

UNIVERSITY OF DEBRECEN
UNIVERSITY AND NATIONAL LIBRARY

List of other publications

4. Guttman, M., **Váradi, C.**, Lee, K.K., Guttman, A.: Comparative glycoprofiling of HIV gp120 immunogens by capillary electrophoresis and MALDI mass spectrometry.
   Electrophoresis. Epub ahead of print (2015)
   DOI: http://dx.doi.org/10.1002/elps.201500054
   IF:3.161 (2013)

5. Szekrényes, Á., Partyka, J., **Váradi, C.**, Krenkova, J., Foret, F., Guttman, A.: Sample preparation for N-glycosylation analysis of therapeutic monoclonal antibodies by electrophoresis.
   In: Microchip Capillary Electrophoresis Protocols : Methods in Molecular Biology. Ann Van Schepdael, Springer protocols, Clifton, N. J, 183-195, 2015.
   DOI: http://dx.doi.org/10.1007/978-1-4939-2353-3_16

6. Meskó, B., Póliska, S., Szamosi, S., Szekanecz, Z., Podani, J., **Váradi, C.**, Guttman, A., Nagy, L.: Peripheral blood gene expression and IgG glycosylation profiles as markers of tocilizumab treatment in rheumatoid arthritis.
   J. Rheumatol. 39 (5), 916-928, 2012.
   DOI: http://dx.doi.org/10.3899/jrheum.110961
   IF:3.258

Total IF of journals (all publications): 18,566
Total IF of journals (publications related to the dissertation): 12,147

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

29 April, 2015