Targeted Biomarker Discovery by High Throughput Glycosylation Profiling of Human Plasma Alpha1-Antitrypsin and Immunoglobulin A

L. Renee Ruhaak1*, Carolien A. M. Koeleman1, Hae-Won Uh2, Jord C. Stam3, Diana van Heemst4, Andrea B. Maier4, Jeanine J. Houwing-Duistermaat2, Paul J. Hensbergen1, P. Eline Slagboom5,6, André M. Deelder1, Manfred Wuhrer1

1 Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands, 2 Section Medical Statistics, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands, 3 Section Biomolecular Imaging, Department of Biology, Utrecht University, Utrecht, The Netherlands, 4 Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands, 5 Section Molecular Epidemiology, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands, 6 Netherlands Consortium of Healthy Aging, Leiden, The Netherlands

Abstract

Protein N-glycosylation patterns are known to show vast genetic as well as physiological and pathological variation and represent a large pool of potential biomarkers. Large-scale studies are needed for the identification and validation of biomarkers, and the analytical techniques required have recently been developed. Such methods have up to now mainly been applied to complex mixtures of glycoproteins in biofluids (e.g. plasma). Here, we analyzed N-glycosylation profiles of alpha1-antitrypsin (AAT) and immunoglobulin A (IgA) enriched fractions by 96-well microtitration plate based high-throughput immuno-affinity capturing and N-glycan analysis using multiplexed capillary gel electrophoresis with laser-induced fluorescence detection (CGE-LIF). Human plasma samples were from the Leiden Longevity Study comprising 2415 participants of different chronological and biological ages. Glycosylation patterns of AAT enriched fractions were found to be associated with chronological (calendar) age and they differed between females and males. Moreover, several glycans in the AAT enriched fraction were associated with physiological parameters marking cardiovascular and metabolic diseases. Pronounced differences were found between males and females in the glycosylation profiles of IgA enriched fractions. Our results demonstrate that large scale immuno-affinity capturing of proteins from human plasma using a bead-based method combined with high-throughput N-glycan analysis is a powerful tool for the discovery of glycosylation-based biomarker candidates.

Introduction

Glycosylation occurs on more than 50% of human proteins [1] and proper glycosylation is essential for the survival of most multicellular organisms. N-Glycans have important functions in several biological processes such as cellular interaction, differentiation, and immunological defense mechanisms [1–6]. Protein N-glycosylation is very diverse, and more than 100 different N-glycans may be found on a specific protein [7]. The glycan signature, i.e. the total plasma N-glycome of an individual, reflects both genetic and physiological factors [8] and is highly reproducible in a given physiological state [9,10]. However, when the physiological state changes, e.g. due to aging or disease, the glycan pattern can change dramatically [9]. Due to the large variability of plasma N-glycans may be found on specific proteins. The glycan signature, i.e. the total plasma N-glycome of an individual, reflects both genetic and physiological factors [8] and is highly reproducible in a given physiological state [9,10]. However, when the physiological state changes, e.g. due to aging or disease, the glycan pattern can change dramatically [9]. Due to the large variability of protein glycosylation and its reflection of physiological and pathological conditions, protein glycosylation patterns have been marked as a large field of potential biomarkers [11].

To identify and validate biomarkers, large-scale studies are needed, and the analytical methods required for the evaluation of protein glycosylation patterns at the glycan level in larger sample sets have only recently been developed (e.g. [12–14]). Up to now, these methods have mainly been applied for the evaluation of plasma N-glycosylation profiles (e.g. [9,15–17]). However, some constraints are associated with the interpretation of the results from such analyses. As the profiles originate from the total protein pool present in plasma, changes in the N-glycosylation profile may be caused by alterations in protein concentration, or by changes in protein-specific glycosylation patterns. Another aspect is that glycans from high abundant proteins dominate the glycan pattern, and changes in the glycosylation signature of less abundant proteins will not be detectable.

There is, therefore, a need for fast, large-scale glycan analysis of specific glycoproteins or groups of glycoproteins. Current methods for glycoprotein enrichment often comprise lectins or (immuno-)affinity purification (e.g. [10,19]). Except for affinity capturing of
IgG [20–22], however, this has to our knowledge not been applied in large-scale studies. Whilst the IgG glycosylation profiling studies applied Protein A and/or Protein G purification, we used in large-scale studies. Whilst the IgG glycosylation profiling studies IgG [20–22], however, this has to our knowledge not been applied in large-scale studies. Whilst the IgG glycosylation profiling studies IgG [20–22], however, this has to our knowledge not been applied in large-scale studies. Whilst the IgG glycosylation profiling studies IgG [20–22], however, this has to our knowledge not been applied in large-scale studies. Whilst the IgG glycosylation profiling studies 

Experimental Section

Materials

Dimethylsulfoxide (DMSO), Nonidet P-40 (NP-40), formic acid, triethylamine (TEA), aminopropene-1,3,6-trisulfonic acid (APTS), 2-picolylamine, trithiocetic acid (TFA), and agarose-linked goat anti-human IgA antibody were obtained from Sigma-Aldrich (Zwijndrecht, The Netherlands). Sodium dodecyl sulfate (SDS) was bought from United States Biochemicals (Cleveland, OH). Porcine Gase F was obtained from Roche Diagnostics (Mannheim, Germany). Biogel P-10 was obtained from Bio-Rad (Veenendaal, The Netherlands) and citric acid from Merck (Darmstadt, Germany). Acetonitrile (ACN) was purchased from Biosolve ( Valkenswaard, The Netherlands). 0.45 μm GHP filter plates were obtained from Pall Corporation (Ann Arbor, MI) and 96-well V-bottom deep well plates from Westburg (Leusden, The Netherlands). PCR plates for measurement in the DNA-sequencer were obtained from Thermo Fisher Scientific via Westburg (Leusden, The Netherlands). 96-well polypyrrole filter plates containing a 10 μm PE frit were obtained from Orochem (Lemhard, IL). Milli Q water was used throughout (Merck Millipore). Standard plasma for the repeatability study and plate-standards was obtained from a healthy donor.

Participants of the Leiden Longevity Study

In the Leiden Longevity Study (LLS), Caucasian families were recruited if at least two long-lived siblings were alive and fulfilled the age-criterion of 89 years or older for males and 91 year or older for females, representing less than 0.1% of the Dutch population in 2001. As no proper controls exist for this age group, for further studies the offspring of these long-lived nonagenarians were included with their partners as controls. The offspring carry on average 50% of the genetic propensity of their long-lived parent and were shown to have a 30% lower mortality rate compared with their birth cohort [36]. Their similarly aged partners, with whom most have had a relationship for decades, were included as age- and environmentally matched controls. In total, 943 long-lived siblings were included, 1671 offspring with a mean age of 59.4 and 744 partners with a mean age of 38.7 years [36]. For the offspring of long-lived siblings and their partners, non-fasting serum samples were taken at baseline for the determination of endocrine and metabolic parameters. Additional information was also obtained from the generation of offspring and controls, including self-reported information on height and weight and information on medical history from the participants’ treating physicians (for details see Table 1 of [15]).

Ethics statement

The study protocol was approved by the Leiden University Medical Center ethical committee and an informed consent was signed by all participants prior to participation in the study.

Phenotypic parameters

All serum measurements were performed with fully automated procedures. For insulin, the Immulite 2500 from DPC (Los Angeles, CA, USA) was applied. For glucose, hCRP, total cholesterol, HDL-cholesterol (HDL-C) and triglycerides the Hitachi Modular or the Cobas Integra 800, both from Roche, Almere, The Netherlands were applied. LDL-cholesterol level (LDL-C) was calculated using the Friedewald formula (LDL-C = total cholesterol – HDL-C – [triglycerides/2.2]); unit mmol/L) and set to missing if plasma triglyceride concentration exceeded 4.52 mmol/L.
**Preparation of anti-AAT beads**

The anti-human AAT llama antibody fragments (VHHs) were kindly provided by BAC (Leiden, The Netherlands). VHHs were purified, dialyzed against PBS and coupled to NHS-activated Sepharose 4 fast flow beads (GE Healthcare) according to the manufacturer’s instructions, at 3 mg/mL overnight. Using SDS-PAGE with Coomassie staining it was confirmed that 80% of the VHHs were coupled to the beads. Beads were stored in 20% PAGE with Coomassie staining.

**Immuno-affinity enrichment of AAT and IgA**

Antibody coated beads were washed three times with 10 volumes of PBS. Six µL of anti-AAT or 20 µL of anti-IgA beads were applied per well of a 96-well filter plate. The volume was brought to 200 µL with PBS, and 10 µL of plasma were applied per well. To each 96-well plate, 2 control samples, 92 study samples, and two blanks were applied. The plate was sealed with tape and incubated on a shaker for 1 h. The beads were washed with 3 x 200 µL PBS by vacuum filtration. After subsequent washing with 2 x 200 µL water, enriched glycoproteins were eluted with 100 µL of 100 mM formic acid into a V-bottom microtitration plate. Samples were dried by vacuum centrifugation. To allow reuse of the antibody coated beads, beads were washed in the filter plates using 3 x 100 µL of 100 mM formic acid, followed by equilibration using 3 x 200 µL PBS. 200 µL of PBS were added to the well prior to application of the following batch of plasma samples.

**Proteomic analysis of anti-AAT and anti-IgA captured proteins**

From three different subjects out of the total study cohort, anti-AAT and anti-IgA captured proteins were dried by vacuum centrifugation, dissolved in 25 µL of 25 mM NH₄HCO₃ and analyzed for their protein constituents using two different approaches.

For the first approach, 10 µL of the reconstituted samples were used for in-solution trypsin digestion. For this purpose, proteins were first reduced and alkylated using diethyldithiol (5 mM in 25 mM NH₄HCO₃, 30 min, 60°C) followed by iodoacetamide (15 mM in 25 mM NH₄HCO₃, 30 min, RT, in the dark). The alkylation reaction was stopped by adding 15 mM of diethylether. Then, 50 ng trypsin (sequence grade modified Trypsin, Promega, Madison, WI) was added and proteins were digested overnight at 37°C. The digestion reaction was stopped using 2 µL of 5% TFA.

The digests were then analyzed by LC-MS using a Ultimate 3000 RSLCnano LC system (Thermo Scientific, Sunnyvale, CA) coupled to a HCTriple ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany). The samples were injected onto an Acclaim C₁₈ PepMap100 trapping column (100 µm x 2 cm, 5 µm, 100 Å (Thermo Scientific) and washed with 100% A (3% ACN in 0.1% formic acid) to 5 µL/min for 6 min. Following valve switching, peptides were separated on an Acclaim C₁₈ PepMap RSLC column (75 µm x 150 mm, 2 µm, 100 Å) at a constant flow of 300 nL/min. The peptide elution gradient was then increased to 40% B (95% ACN in 0.1% formic acid) in 48 min followed by an increase to 65% B in 10 min. The nanoflow LC was coupled to the mass spectrometer using a nano-electrospray ionization source. The spray voltage was set at 1.2 kV and the temperature of the heated capillary was set to 165°C. Eluting peptides were analyzed using the data dependent MS/MS mode over a 300-1500 m/z range. The five most abundant ions in an MS spectrum were selected for MS/MS analysis by collision-induced dissociation using helium as the collision gas.

Peak lists were generated using DataAnalysis 4.0 software (Bruker Daltonics) and exported as Mascot Generic (MGF) files. These files were searched against the human Swissprot database (Date of release 21-05-2012, 536029 sequences, 20319 Homo sapiens) with the Mascot search algorithm (Mascot 2.2, Matrix Science, London, UK) using Mascot Dcamon. An MS tolerance of 0.6 Da (with        13C = 1) and a MS/MS tolerance of 0.5 Da was used. Trypsin was designated as the enzyme and up to two missed cleavage sites was allowed. Carbamidomethylcysteine was selected as a fixed modification and oxidation of methionine as a variable modification. Only significant protein hits with at least two unique peptides and a score above 30 were selected.

For the second approach, anti-AAT and anti-IgA captured proteins were analyzed by SDS-PAGE using a pre-cast gradient gel [Invitrogen NuPAGE 4–12% Bis-Tris Gel] and proteins were visualized using Colloidal Blue Staining (Life Technologies). The major protein band from the anti-AAT and anti-IgA samples was selected for in-gel digestion. For this purpose, bands were cut into small pieces, and washed with 25 mM NH₄HCO₃ followed by dehydration with ACN for 10 min. For reduction and alkylation, dried gel particles were first incubated with 10 mM diethyldithiotreitol for 30 min at 56°C. Following dehydration with ACN, gel plugs were subsequently incubated in 55 mM iodoacetamide for 20 min
at room temperature. After two rounds of washing with 25 mM NH₄HCO₃ and dehydration with ACN, the gel particles were completely dried by vacuum centrifugation. The dried gel particles were re-swollen for 15 min on ice after the addition of 15 μL of a trypsin solution (5 ng/μL in 25 mM NH₄HCO₃). Following this, 20 μL of 25 mM NH₄HCO₃ was added and the samples were kept on ice for an additional 30 min. Tryptic digestion was subsequently performed overnight at 37°C. The overlaying digestion-solution containing the tryptic peptides was collected (extract 1). One additional round of extraction with 20 μL 0.1% TFA was used to extract peptides from the gel plugs and this extract was pooled with extract 1.

MALDI-ToF-MS analyses were performed on an Ultraflex II mass spectrometer (Bruker Daltonics) using 2,5-dihydroxybenzoic acid (5 mg/mL of 50% ACN/0.1% trifluoroacetic acid) as a matrix. The mass spectrometer was used in the positive ion reflector ion mode. Spectra were imported in FlexAnalysis 3.0 (Bruker Daltonics) for smoothing, baseline subtraction and peak picking. Peak lists were searched against the human Swissprot database using the Mascot search algorithm in Biotools (version 3.2, build 3.34, Bruker Daltonics). Trypsin was selected as the enzyme and one missed cleavage was allowed. Carbamidomethylation cysteine was selected as a fixed modification and oxidation of methionine was variable. The MS tolerance was set at 100 ppm.

Glycan release, labeling and purification

N-glycans were released from the enriched glycoprotein samples similarly to the procedure described previously for plasma [14] with slight modifications. Shortly, proteins were denatured after resuspension in 2 μL 2% SDS by incubation at 60°C for 10 min. Subsequently, 2 μL 2% NP-40 containing 0.2 mM of PNGase F was added to the samples. The samples were incubated overnight at 37°C for N-glycan labeling. Release of oligosaccharides was performed as published [35,37] with slight modifications: 2 μL of a freshly prepared solution of label (APTS; 20 mM in 3.6 M citric acid) and 2 μL of freshly prepared reducing agent solution (0.2 M 2-picoline borane in DMSO, see [38]) were added, the plate was sealed using adhesive tape and after 5 min of shaking, the samples were incubated at 37°C for 16 h.

Hydrophilic interaction liquid chromatography (HILIC)-SPE was applied to purify labeled N-glycans. An amount of 100 μL of a 100 mg/mL suspension of Biogel P-10 in water/ethanol/ACN (70:20:10, v/v) was applied to each well of a 0.45 μm GHP filter plate (Pall Corporation, Ann Arbor, MI). Solvent was removed by application of vacuum using a vacuum manifold (Millipore, Billerica, MA). All wells were prewashed using 5×200 μL water, followed by equilibration using 3×200 μL ACN/water (80:20, v/v). The samples were loaded to the wells, and the plate was shaken for 5 min. The wells were subsequently washed using 3×200 μL ACN/100 mM triethylamine (TEA) adjusted to pH 8.5 with acetic acid (80:20, v/v), followed by 3×200 μL ACN/water (80:20, v/v). Washing steps were performed by addition of solutions, incubation for 30 s, and removal of solvent by vacuum. Water (100 μL) was applied followed by 5 min incubation to allow swelling of the Biogel P-10 particles. For elution, samples were incubated for 5 min with 200 μL water prior to collection of eluates by vacuum in a 96-well V-bottom polypropylene deep well plate. The elution step was repeated, and the 2×200 μL eluates were combined and analyzed immediately by CGE-LIF or stored at −20°C until usage.
glucose as well as insulin level (ln transformed) – adjusting for the family status (offspring or partner of the offspring) age, sex and the age × sex interaction.

To determine potential biomarkers for longevity, logistic regression was applied to investigate whether glycosylation feature (independent variable) was predictive in classifying the family status after adjustment for age, sex, and their interaction. In that respect, the response variable (the family status) is coded as 0 (= partner) and 1 (= offspring of long-lived sibling). Finally, linear regression was used to explore relationships between glycosylation features and disease status – levels of CRP and prevalence of MI, cardiovascular diseases or diabetes- adjusting for age, sex and their interaction.
Protein enrichment

Results

Protein enrichment

Glycoproteins were enriched from human plasma using high-throughput immuno-affinity capturing in a 96-well format. Anti-AAT and anti-IgA antibody coated beads were applied to each of the wells of a filter plate, after which plasma was added to the wells. After incubation for 1 h and thorough washing with PBS and water, the enriched glycoprotein fraction was eluted using 100 mM formic acid.

To check the purity of the enriched glycoprotein samples, we performed a proteomic analysis of the samples from three different subjects from our study cohort. First, we analyzed these samples by SDS-PAGE. In addition, we performed an in-solution tryptic digestion of the enriched fraction. The digests were subsequently analyzed by MALDI-ToF-MS and the resulting peptide mass fingerprints were searched against the human Swissprot database. As expected, this resulted in the identification of AAT and IgA1 (heavy chain), respectively and the major ion trap-MS analysis (Tables ST1 and ST2 in Supporting Information S1).

We recently developed a platform for high-throughput analysis of APTS-labeled glycans using a multiplexed CGE-LIF instrument [35]. Using this platform, 96 samples can be measured within 3 hours, which makes it highly suitable for large-scale studies. Previously, we obtained annotations of 35 abundant plasma N-glycan structures in CGE-LIF by employing glycan standards whenever available and elucidating additional structures by HILIC-HPLC and MALDI-Fourier transform ion cyclotron resonance-MS. This previous annotation of plasma N-glycans, together with literature, was now used to annotate the observed APTS-labeled glycans in the AAT- and IgA-enriched samples. An overview of the annotation is given in Table 1 (AAT) and Table 2 (IgA). An overlay of ten annotated electropherograms that were randomly chosen from the LLS samples is depicted in Figure 1.

Repeatability of the method

To assess the repeatability of the procedure for large-scale protein enrichment and subsequent glycan analysis, the intra-batch as well as the inter-batch variations were calculated. The intra-batch relative standard deviation (RSD) for the AAT enriched fraction averaged over the 21 peaks was found to be 15.8% (peak 3 not included as it was used for normalization). The inter-batch RSD for the IgA enriched fraction averaged over 14 peaks was found to be 20.2% (peak 5 not included as it was used for normalization). The inter-batch standard deviation was determined by comparing the results from the four experiments, these samples (Fig. SF1 in Supporting Information S1). In two of the samples (from subject 2 and 3) we also identified IgA2 which is known to be the minor IgA species within the circulation [25]. In one of the samples (subject 1), 2 peptides from IgG were observed indicating trace amounts of IgG, which will probably not affect the results from our IgA glycan profiling analysis.

Table 2. Peak annotation of IgA-enriched glycans.

| Fraction | Composition | Spiking | Literature | PGC-LC-ESI-IT-MS |
|----------|-------------|---------|------------|------------------|
| IgA-1    | Hex5HexNAc5NeuAc2 | X | [29,30] | 1111.9 [M+H]2+ |
| IgA-4    | Hex5HexNAc5NeuAc2 | X | [29,30] | 1184.9 [M+H]2+ |
| IgA-5    | Hex5HexNAc4Fuc1NeuAc2 | X | [29,30] | 1286.4 [M+H]2+ |
| IgA-6    | Hex5HexNAc4Fuc1NeuAc2 | X | [29,30] | 966.4 [M+2H]2+ |
| IgA-7    | Hex5HexNAc4Fuc1NeuAc2 | X | [29,30] | 1039.4 [M+2H]2+ |
| IgA-8    | Hex5HexNAc4Fuc1NeuAc2 | X | [29,30] | 1140.9 [M+2H]2+ |
| IgA-9    | Hex5HexNAc4Fuc1NeuAc2 | X | [29,30] | 1184.9 [M+H]2+ |
| IgA-10   | Hex5HexNAc5Fuc1NeuAc2 | X | [29,30] | 966.4 [M+2H]2+ |
| IgA-11   | Hex4HexNAc5Fuc1 | X |            |                  |
| IgA-14   | Hex5HexNAc4Fuc1 | X |            |                  |
| IgA-15   | Hex5HexNAc5Fuc1 | X |            |                  |

Glycan compositions were annotated based on spiking experiments with human plasma and literature. Compositions are given in terms of Hex: hexose, HexNAc: N-acetylhexosamine, Fuc: fucose, NeuAc: N-acetylneuraminic acid. doi:10.1371/journal.pone.0073082.t002

Analyses were performed using STATA 10 (StataCorp LP, College Station, Texas, USA) and R version 2.9.0 (R Development Core Team).

Figure 2. Repeatability of the enrichment procedure with subsequent analysis. Relative intensities together with their inter-batch standard deviations are depicted for AAT (A) and IgA (B). doi:10.1371/journal.pone.0073082.g002

Glycan peak annotation

We recently developed a platform for high-throughput analysis of APTS-labeled glycans using a multiplexed CGE-LIF instrument [35]. Using this platform, 96 samples can be measured within 3 hours, which makes it highly suitable for large-scale studies. Previously, we obtained annotations of 35 abundant plasma N-glycan structures in CGE-LIF by employing glycan standards whenever available and elucidating additional structures by HILIC-HPLC and MALDI-Fourier transform ion cyclotron resonance-MS. This previous annotation of plasma N-glycans, together with literature, was now used to annotate the observed APTS-labeled glycans in the AAT- and IgA-enriched samples. An overview of the annotation is given in Table 1 (AAT) and Table 2 (IgA). An overlay of ten annotated electropherograms that were randomly chosen from the LLS samples is depicted in Figure 1.
and is depicted in Figure 2. The average inter-batch RSD for the 21 peaks observed in the AAT enriched fraction was found to be 33.8%. The average inter-batch RSD for the 14 peaks in the IgA enriched fraction was found to be 32.1%.

Relation of protein glycosylation to chronological age and sex

Glycosylation profiles of the AAT enriched fraction could be obtained from 2271 individuals in the LLS, while profiles from the IgA enriched fraction were obtained from 2382 individuals. The participants in the LLS cover a large age-range (30–80 years). After batch-correction, the effects of age and sex were evaluated in both datasets (see Fig. 3, Tables ST3 and ST4 in Supporting Information S1). Several effects can be observed: non-fucosylated triantennary structures in the AAT enriched fraction (AAT-2, AAT-8, AAT-9 and AAT-16) are all negatively correlated to chronological age. Moreover, levels of most glycan features found in the IgA enriched fraction are less abundant in males than in females. Previous studies have reported that chronological age-associated changes in plasma glycosylation are usually more profound in females than in males [16,21,41]. Therefore, we investigated whether the interaction between chronological age and sex influences the glycosylation patterns of the AAT and IgA enriched fractions, but no significant interaction was found (data not shown).

Protein glycosylation is affected by physiological parameters

In two recent studies, plasma glycosylation patterns were found to be associated with altered lipid status and changes in BMI and glucose homeostasis [15,16]. We evaluated whether glycosylation features of the AAT and IgA enriched fractions were associated with BMI, plasma levels of cholesterol, HDL-cholesterol, LDL-cholesterol, triglycerides, glucose and insulin activity. While no associations could be observed in the IgA enriched fraction (data not shown), levels of HDL-cholesterol, BMI, insulin levels, and especially levels of triglycerides are associated with several glycan features observed in the AAT enriched fractions (see Fig. 4 and Table ST5 in Supporting Information S1). This immediately raises the question whether the changes in the glycosylation profile of the AAT enriched fraction may be associated with health or disease states in general and, therefore, could be linked to longevity.

Relation of protein glycosylation with familial longevity

The Leiden Longevity study was originally designed to study biological mechanisms that influence human life span. Here we evaluated whether glycosylation patterns of AAT and/or IgA enriched fractions are associated with familial longevity. Using a logistic regression model, none of the glycosylation features showed a significant difference between the group of offspring and the group of partners (data not shown). In addition, due to the association of various glycosylation features in the IgA- and AAT enriched fractions with sex we tested for association with familial longevity after sex stratification. Again, no significant differences were found between offspring and partners (data not shown).

Altered protein glycosylation is associated with disease

While the glycosylation features could not be linked to familial longevity, associations were found between glycosylation patterns of the AAT enriched fractions and several physiological parameters (Fig. 4 and Table ST5 in Supporting Information S1). This raised the question whether glycan in the AAT enriched fractions may reflect specific disease states. High levels of plasma C-reactive protein (CRP) indicate an acute inflammatory response, and may be regarded as a marker for the presence of inflammatory diseases [42]. Moreover, plasma N-glycosylation was previously reported to be correlated to CRP levels [43]. Therefore, we assessed the relation between CRP levels and glycosylation features of the AAT enriched fractions. The results are depicted in Figure 4 and Table ST6 in Supporting Information S1. A large number of glycosylation features were found to be associated with CRP levels with high significance, indicating that the glycosylation profiles of AAT enriched fractions are indeed altered in diseased individuals showing systemic inflammation.
Discussion

We here present the first study on large scale protein enrichment using immuno-affinity capturing with subsequent high-throughput N-glycan analysis. In the present study we have chosen to use a bead-based enrichment strategy, using commercially available anti-IgA antibody coated beads and anti-AAT Llama antibody coated beads. We also decided to use – for the first time in biomarker discovery – a real high-throughput analytical glycan profiling technique (CGE-LIF), which can be performed with broadly available instrumentation (Sanger DNA sequencer). Multiplexed CGE-LIF technology has great potential in biomarker discovery as well as product monitoring in biotechnological facilities due to the very fast analysis times, high resolution power, easy operating procedures and capability to separate N-glycan isomers. Alternative strategies include mass spectrometry using MALDI-MS (e.g. [44,45]) or direct infusion ESI-MS, which require substantial training on the instrument and do not provide isomer separation, and UPLC with fluorescence detection (e.g. [22]), which provides isomer separation but has much lower throughput as it is not multiplexed.

The intra-batch repeatability of the strategy described in this study featured an average RSD of 16% and 20% for the AAT and IgA enriched fractions, respectively. This value includes the variation due to the protein enrichment, glycan release, labeling, purification and measurement. We have previously shown that the contribution of the measurement to this variation is rather small (6.2% [35]). The large inter-batch variation (average RSDs of approximately 33%) makes batch-corrections necessary when analyzing large sample sets.

Recent studies in human plasma have revealed that chronological age and sex influence the plasma N-glycosylation pattern [15,16]. Likewise, our large scale enrichment of AAT and IgA from plasma samples from participants of the LLS using antibody affinity capturing followed by subsequent high-throughput N-glycan analysis, revealed that glycosylation patterns of the AAT enriched fractions are associated with chronological age and differ between females and males (Figure 3 and Table ST3 in Supporting Information S1). Interestingly, for the AAT enriched fraction, most of the glycans that are decreased with increasing chronological age are non-fucosylated sialylated compounds, while one of the fucosylated and sialylated glycans (AAT-11) is positively related to age. Levels of several oligosaccharides from the AAT enriched fractions are higher in females, as depicted by their negative correlation with sex. Two triantennary structures with fucose are higher in males, while the non-fucosylated triantennary structures are more abundant in females.

Figure 4. N-glycans on AAT are associated with physiological parameters. Regression coefficients of the linear regression analysis are plotted for N-glycans from AAT for BMI (A), cholesterol (B), HDL-cholesterol (C), LDL-cholesterol (D), triglycerides (E), glucose (F), insulin level (G) and C-reactive protein (H). Error bars indicate standard error and significance is indicated with an asterisk. No associations were obtained for Variable AAT-3, since the data was normalized to an individuals’ level of this peak.

doi:10.1371/journal.pone.0073082.g004
Glycosylation profiles of the IgA enriched fractions, on the other hand, hardly showed any chronological-age related alterations. Only a mono-sialylated biantennary glycan (IgA-11) could be shown to increase with increasing chronological age. This result is surprising, as it is well known that the level of galactosylated glycan on the most abundant immunoglobulin IgG decreases with chronological age [21,22,41,46,47]. Levels of several di- and non-sialylated biantennary glycans with or without bisecting GlcNAc in the IgA enriched fractions were observed to be lower in males than in females. Similarly, for IgG several glycoforms were observed at lower levels in male individuals [21].

Various differences between offspring of longlived siblings and age-matched partners of the offspring have so far been reported. It could be observed that the offspring has a lower prevalence of myocardial infarction, hypertension and diabetes mellitus [40]. In the non-diabetic individuals, non-fasted serum glucose levels were decreased in the offspring [49]. Moreover, in the offspring, larger LDL-cholesterol particles were reported [50]. We recently reported the association of two whole serum glycosylation features [15] as well as serum IgG glycosymes [21] with familial longevity. Here, we did not find an association of glycosylation features from AAT and IgA enriched fractions with familial longevity.

However, further analysis revealed an association of several glycans in the AAT enriched fractions with parameters marking cardiovascular, metabolic and inflammatory diseases (Figure 4, Tables ST3 and ST6 in Supporting Information S1). Two biantennary structures (AAT-2 and AAT-9) are positively related to both cholesterol and triglyceride levels, but negatively associated with CRP levels. Several biantennary glycans (AAT-4, AAT-6 and AAT-13) were negatively associated with BMI, CRP, triglyceride, glucose and insulin levels, but positively associated with HDL. Other glycans (AAT-1, AAT-5, AAT-7 and AAT-14) were positively associated with BMI, CRP, triglyceride, glucose and insulin levels, but negatively related to HDL. This confirms the negative relation of HDL levels with cardiovascular, metabolic and inflammatory diseases relative to BMI, CRP, triglyceride, glucose and insulin levels. Furthermore, it indicates a relation between the glycosylation of the liver-derived protein AAT with such diseases. While a relation of IgA glycosylation with disease parameters was also investigated, no associations were observed, indicating that B-cell derived proteins are less affected by metabolic diseases than liver cell derived proteins.

The proteomics analysis of the affinity-enriched IgA and AAT samples indicated a good purity of the samples. Using SDS-PAGE of IgA and AAT enriched fractions, it was demonstrated that the target proteins represent the most strongly stained components, but there are also some co-purified species, which are highly likely proteins that are bound to the protein of interest (either AAT or IgA). In-solution proteomics revealed the presence of some additional glycoproteins that would contribute to the N-glycosylation patterns obtained in this study. However, as these additional glycoproteins appeared to be of low abundance, their contribution to the observed glycosylation profiles of the enriched fractions is expected to be minor. This is in line with the fact that the observed glycosylation of IgA and AAT enriched fractions is largely in line with literature (see Tables 1 and 2). While the immunocapturing approach described here is suitable for enrichment and purification of medium to high abundant glycoproteins, other platforms will have to be considered for glycoproteins that occur at even lower abundances in complex mixtures such as human plasma. Such platforms will need to further reduce aspecific binding by minimizing contact surfaces in order to allow efficient affinity-enrichment.

The roles of glycosylation, and particularly N-glycosylation on specific proteins are not very well understood. So far, the functions of the N-glycans on immunoglobulin G are most studied and best understood. However, it may be expected that glycosylation of other plasma proteins is similarly important. In support of this hypothesis, levels of sialylation on EPO influence the proteins activity and half life [51]. Moreover, protein clearance from the blood is mediated by the asialoglycoprotein receptor, which has affinity for terminal β-galactose residues [32].

In our large-scale cross-sectional study various associations were found – similar to IgG – which indicate that actually also these two plasma proteins – IgA and AAT – show a highly regulated glycosylation that is influenced by and/or influencing various physiological parameters. Therefore, our data provide very strong support for the complex regulation of plasma protein glycosylation. Of particular relevance are the differences observed between IgG and IgA glycosylation: while the glycosylation profiles are qualitatively very similar (vast overlap of observed glycan species), there is a huge quantitative difference, with mainly a much lower degree of sialylation for IgG (Fig. SF3 in Supporting Information S1). Intriguingly, these two glycoproteins show very different dependencies on physiological parameters, which indicates the distinct regulation of IgA and IgG glycosylation, whilst they are both produced by plasma cells. As for IgA, it is worth mentioning that the observed CGE-LIF glycan profiles are qualitatively very similar to IgG glycan profiles, yet there are vast quantitative differences. The IgG N-glycosylation pattern has been characterized before using CGE-LIF [53], and even though many of the glycans observed on IgA can also be identified on IgG, their levels differ significantly. To illustrate this, an electropherogram of IgG as well as IgA derived N-glycans is depicted in Fig. SF2 in Supporting Information S1. Clearly, neutral glycans eluting in the latter part of the electropherogram (after 9000 dp) are of very high abundance on IgG, while only minor peaks are observed on IgA. This is in line with our proteomics analyses indicating that IgG contamination of the IgA samples is negligible. Interestingly, the neutral glycans found in the electropherogram of the IgA enriched fraction (IgA-11, IgA-14 and IgA-15) could not be related to chronological age, while the corresponding glycans of IgG were previously shown to be highly associated with chronological age [21,22,41,46,47].

More indications for the differential regulation of IgG and IgA glycosylation can be found in literature: in rheumatoid arthritis, the level of galactosylation of IgG decreases dramatically (e.g. [54–56]), while N-glycans in the IgA enriched fractions are hardly observed at lower levels in male individuals [21].

In conclusion, it was shown that a strategy consisting of large-scale immuno-affinity capturing of proteins from human plasma using a bead-based method, coupled with high-throughput N-glycan analysis using multiplexed CGE-LIF is a powerful tool for the analysis of N-glycosylation patterns of specific glycoproteins in large studies. The strategy has successfully been applied to approximately 2400 plasma samples from the LLS demonstrating for the first time the complex regulation of plasma protein glycosylation for more typical plasma proteins such as IgA and AAT and not just for the IgG molecule with its Fe N-glycans that are only partially accessible.
References

24. Janciauskiene SM, Bals R, Koczulla R, Vogelmeier C, Kohnlein T, et al. (2011) Human plasma glycosylation: a personal perspective. Ann N Y Acad Sci 1248: 46–71.

23. Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 1: 845–867.

21. Ruhaak LR, Uh HW, Beekman M, Koeleman CA, Hokke CH, et al. (2010) Human serum processing and analysis methods for rapid and reproducible N-glycan analysis. J Proteome Res 9: 694–701.

20. van de Geijn FE, Wuhrer M, Selman MH, Willemsen SP, de Man YA, et al. (2010) A high throughput glycoprotein profiling system for analysis of human plasma samples. J Chromatogr A 1217: 216–224.

19. Shikata K, Yasuda T, Takeuchi F, Konishi T, Nakata M, et al. (1998) Structural characterization of the membrane glycoprotein CD59. J Biol Chem 273: 7229–7244.

18. Poland DC, Schalkwijk CG, Stehouwer CD, Koeleman CA, van het HB, et al. (2010) The glycosylation of the complement regulatory protein, human erythrocyte CD59. J Biol Chem 272: 7229–7244.

17. Vanhooren V, Liu XE, Franceschi C, Gao CF, Libert C, et al. (2009) N-glycan profiles as tools in diagnosis of hepatocellular carcinoma and prediction of human longevity. J Proteome Res 10: 1667–1674.

16. Knezevic A, Gornik O, Polasek O, Pucic M, Redzic I, et al. (2010) Effects of dietary saturated fatty acids on human plasma N-glycosylation profiles using correlation optimised warping. J Chromatogr A 805: 17–35.

15. Marin HD, Campbell HM, McPherson J, Florin EL, McFarland SA, et al. (2009) Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. Anal Biochem 389: 112–121.

14. Ruhaak LR, Hokke CH, Beekman M, Stehouwer CD, Koeleman CA, et al. (2010) Human serum processing and analysis methods for rapid and reproducible N-glycan analysis. J Proteome Res 9: 694–701.

13. Royle L, Campbell MP, Radcliffe CM, White DM, Harvey DJ, et al. (2008) Identification of alpha(1)-antitrypsin variants in plasma with the use of proteomic technology. Clin Chem 47: 2012–2022.

12. Laroy W, Contreras R, Callewaert N (2006) Glycome mapping on DNA microarrays. J Proteome Res 5: 1764–1774.

11. Packer NH, von der Lieth CW, Aoki-Kinoshita KF, Lebrilla CB, Paulson JC, et al. (2006) Variability, Heritability and Environmental Determinants of Human Plasma N-Glycome. J Proteome Res 8: 694–701.

10. Rudd PM, Morgan BP, Wormald MR, Harvey DJ, van den Berg CW, et al. (2007) The eGlycoScope Glycan Profiling System: LRR MW HU PJH PES AMD. Contributed reagents/materials/analysis tools: ABM DvH JCS. Wrote the paper: LRR MW HU PJH PES AMD.

9. Gomes MM, Wall SB, Takahashi K, Novak J, Renfrow MB, et al. (2008) Identification of IgA1 N-Glycosylation and its contribution to FcαRIβ-binding. Biochemistry 47: 11267–11275.

8. Mattu TS, Pless RJ, Willis AC, Kilian M, Wormald MR, et al. (1998) The glycosylation and structure of human serum IgA1 Fab, Fc and Fcα regions and the role of N-glycosylation on Fc alpha receptor interactions. J Biol Chem 273: 2260–2272.

7. Rudd PM, Morgan BP, Wormald MR, Harvey DJ, van den Berg CW, et al. (2006) Structure, variability and diagnostic prospects. Mol Cell Proteomics 1: 845–867.

6. Ohtsubo K, Marth JD (2006) Glycosylation in cellular mechanisms of health and disease. Cell 126: 853–867.

5. Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 1: 845–867.

4. Takahashi M, Kuroki Y, Ohtsubo K, Taniguchi N, et al. (2009) Core fucose and bisecting GlcNAc, the direct modifiers of the N-glycan core: their functions and target proteins. Carbohydr Res 344: 1387–1390.

3. Fukuda MN, Sasaki H, Lopez L, Fukuda M (1989) Survival of recombinant erythropoietin in the circulation: the role of carbohydrates. Blood 73: 84–89.

2. Gu J, Sato Y, Kariya Y, Isaji T, Taniguchi N, et al. (2009) A mutual regulation between cell-cell adhesion and N-glycosylation: implication of the bisecting GlcNAc for biological functions. J Proteome Res 8: 431–435.

1. Shrental-Becher D, Levy Y (2009) Folding of glycoproteins: toward understanding the biophysics of the glycosylation code. Curr Opin Struct Biol 19: 524–533.

Acknowledgments

We thank Matthias Borowiak for helping with data processing and Evelyne Steenwoorden and Irina Dragun for technical assistance.

Author Contributions

Conceived and designed the experiments: PES AMD MW. Performed the experiments: LRR CAMK PJH. Analyzed the data: LRR HU JJH. Contributed reagents/materials/analysis tools: ABM DvH JCS. Wrote the paper: LRR MW HU PJH PES AMD.
49. Rozen MP, Westendorp RG, Frolich M, de Craen AJ, Beckman M, et al. (2009) Human insulin/IGF-1 and familial longevity at middle age. Aging (AlbanyNY) 1: 714–722.
50. Heijmans BT, Beekman M, Hosking-Duistermaat JJ, Cobain MR, Powell J, et al. (2006) Lipoprotein particle profiles mark familial and sporadic human longevity. PLoS Med 3: e495.
51. Su D, Zhao H, Xia H (2010) Glycosylation-modified erythropoietin with improved half-life and biological activity. Int J Hematol 91: 230–244.
52. Sorensen AL, Rumjanneva V, Nayeb-Hashemi S, Clausen H, Hartvig JH, et al. (2009) Role of sialic acid for platelet life span: exposure of beta-galactose residues in the rapid clearance of platelets from the circulation by asialoglycoprotein receptor-expressing liver macrophages and hepatocytes. Blood 114: 1645–1654.
53. Huhn C, Rabaak LH, Wührer M, Beeldner AM (2012) Hexapeptide library as a universal tool for sample preparation in protein glycosylation analysis. J Proteomics 75: 1515–1528.
54. Wada Y, Tajiri M, Ohshima S (2010) Quantitation of saccharide compositions of O-glycans by mass spectrometry of glycopeptides and its application to rheumatoid arthritis. J Proteome Res 9: 1367–1373.
55. Parekh RB, Dwek RA, Sutton BJ, Fernandes DL, Leung A, et al. (1985) Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. Nature 316: 452–457.
56. van Zeben D, Rook GA, Hazes JM, Zwinderman AH, Zhang Y, et al. (1994) Early agalactosylation of IgG is associated with a more progressive disease course in patients with rheumatoid arthritis: results of a follow-up study. Br J Rheumatol 33: 36–43.
57. Field MC, matayakul-Chanler S, Rademacher TW, Rudd PM, Dwek RA (1994) Structural analysis of the N-glycans from human immunoglobulin A1: comparison of normal human serum immunoglobulin A1 with that isolated from patients with rheumatoid arthritis. Biochem J 299 (Pt 1): 261–273.
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Ruhaak, LR; Koeleman, CAM; Uh, H-W; Stam, JC; van Heemst, D; Maier, AB; Houwing-Duistermaat, JJ; Hensbergen, PJ; Slagboom, PE; Deelder, AM; Wuhrer, M

Title:
Targeted Biomarker Discovery by High Throughput Glycosylation Profiling of Human Plasma Alpha1-Antitrypsin and Immunoglobulin A

Date:
2013-09-09

Citation:
Ruhaak, L. R., Koeleman, C. A. M., Uh, H. -W., Stam, J. C., van Heemst, D., Maier, A. B., Houwing-Duistermaat, J. J., Hensbergen, P. J., Slagboom, P. E., Deelder, A. M. & Wuhrer, M. (2013). Targeted Biomarker Discovery by High Throughput Glycosylation Profiling of Human Plasma Alpha1-Antitrypsin and Immunoglobulin A. PLOS ONE, 8 (9), https://doi.org/10.1371/journal.pone.0073082.

Persistent Link:
http://hdl.handle.net/11343/198083

File Description:
Published version

License:
CC BY