Increased Complexity of Plasma N-glycome is Associated with Higher Risk of Type 2 Diabetes Development

Toma Keser[1], Ivan Gornik[2], Frano Vučković[3], Najda Selak[1], Tamara Pavić[1], Edita Lukić[2], Ivan Gudelj[3], Hrvoje Gašparović[4], Bojan Biočina[4], Therese Tilin[5], Annika Wennerström[6], Satu Männistö[6], Veikko Salomaa[6], Aki Havulinna[6], Wei Wang[7,8], James F. Wilson[9], Nishi Charutvedi[5], Markus Perola[6], Harry Campbell[9], Gordan Lauc[1,3] and Olga Gornik[1]

1 University of Zagreb Faculty of Pharmacy and Biochemistry, Zagreb, Croatia
2 Clinical Hospital Centre Zagreb, Zagreb, Croatia
3 Genos Glycoscience Research Laboratory, Zagreb, Croatia
4 Department of Cardiac Surgery, University Hospital Center Zagreb, University of Zagreb School of Medicine, Zagreb, Croatia
5 University College London, Institute of Cardiovascular Science, London, UK
6 National Institute for Health and Welfare (THL), Helsinki, Finland
7 School of Medical and Health Sciences, Edith Cowan University, Perth, Australia
8 Beijing Municipal Key Laboratory of Clinical Epidemiology, School of Public Health, Capital Medical University, Beijing, China
9 Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, UK

Corresponding author: Olga Gornik, University of Zagreb Faculty of Pharmacy and Biochemistry, Ante Kovačića 1, 10 000 Zagreb, Croatia, email: ogornik@pharma.hr, phone: +385 1 6394 449, fax: + 385 1 6394 400

Word count:

Abstract: 258

Main text: 3687

Abstract:

Aims/Hypothesis: Better understanding of type 2 diabetes and its prevention is a pressing need. Changes in human plasma N-glycome are associated with many diseases and represent promising diagnostic and prognostic biomarkers. Variations in glucose metabolism directly affect glycosylation through the hexosamine pathway, but studies of plasma glycome in type 2 diabetes are scarce.

Methods: Using chromatographic approach, we analysed N-linked glycans from plasma proteins in two populations comprising individuals with registered hyperglycaemia during critical illness (increased risk for development of type 2 diabetes) and individuals who stayed normoglycaemic during the same condition - AcuteInflammation (59 cases vs. 49 controls) and AcuteInflammation Replication (52 cases vs. 14 controls) populations. Additionally, N-glycome was also studied in individuals from FinRisk (37 incident cases of type 2 diabetes collected at baseline vs 37 controls), ORCADES (94 individuals with HbA_{1c} > 6.5 % (47.5 mmol/mol) vs 658 controls) and SABRE cohort studies (307 individuals with HbA_{1c} > 6.5 % (47.5 mmol/mol) vs 307 controls).

Results: Individuals with increased risk for diabetes type 2 development (AcuteInflammation and AcuteInflammation Replication populations), incident cases of type 2 diabetes collected at baseline (FinRisk population) and individuals with elevated HbA_{1c} (ORCADES and SABRE populations) all presented increased branching, galactosylation and sialylation of plasma protein N-glycans and these changes were of similar magnitude.

Conclusions/Interpretation: Increased complexity of plasma N-glycan structures is associated with higher risk of developing type 2 diabetes and poorer regulation of blood glucose levels.

Although further research is needed, this finding could offer a potential new approach for improvement in prevention of diabetes and its complications.

Keywords: diabetes predisposition; hyperglycaemia; N-linked glycans; plasma N-glycome; type 2 diabetes;

Abbreviations

2-AB - 2-aminobenzamide; ACN - acetonitrile; AI - AcuteInflammation; AI_R - AcuteInflammation Replication; BEH - bridged ethylene hybrid; BMI - Body Mass Index; CI - confidence interval; DMSO - dimethyl sulfoxide; FIN - FinRisk; G0 - agalactosylated; G1 - monogalactosylated; G2 - digalactosylated; G3 - trigalactosylated; G4 - tetragalactosylated; GlcNAc - N-acetylglucosamine; HB - high branching; HG - hyperglycaemia; HILIC - hydrophilic interaction liquid chromatography; ICU - intensive care unit; LB - low branching; OGTT - oral glucose tolerance test; ORK - ORKNEY; PNGase F - N-glycosidase F; RR - relative risk; S0 - not sialylated; S1 - monosialylated; S2 - disialylated; S3 - trisialylated, S4 - tetrasialylated; SAB - SABRE; SDS - sodium dodecyl sulfate; SE - standard error; UDP - uridine diphosphate; UPLC - ultra performance liquid chromatography;

Introduction

Type 2 diabetes represents one of the major public health challenges globally, having a constantly growing prevalence with the global projection of 7.3 % for the year 2025 [1]. Intensive research in the past decades resulted in the introduction of new oral and parenteral medications that may improve the management of the disease. However, recognising individuals with increased risk of type 2 diabetes development is of equal importance, since many potential interventions are available which can postpone and maybe even prevent the disease onset.

It has been recently shown that individuals without a history of diabetes, who developed hyperglycaemia (plasma glucose > 7.7 mmol/L) during critical illness are at significantly increased risk of developing type 2 diabetes [2, 3]. The relative risk (RR) of type 2 diabetes development during five years after the acute illness was 5.6 (95% confidence interval (CI) 3.1, 10.2). The phenomenon probably arises from the fact that people who develop hyperglycaemia in critical illness already have a latent disorder in glycaemic control. This disorder cannot be detected with the usual screening tests, but manifests in the circumstances of severe acute illness due to inflammatory and stress mediators. Individuals without this latent disorder manage to overcome hyperglycaemic effects of stress, and inflammation and remain normoglycaemic. After the acute illness subsides, individuals who presented with hyperglycaemia become normoglycaemic again, but the underlying disorder in glucose metabolism apparently makes them prone to type 2 diabetes development.

Glycosylation is the most frequent co- and post-translational modification of proteins, which in many cases modulates their function [4]. Contrary to glycation (which is a non-enzymatic chemical reaction), this highly complex enzymatic process is strictly regulated by the network of many enzymes, transcriptional factors, sugar nucleotides and other molecules [5].

Although proteins possess well defined glycosylation sites, a great heterogeneity of their glycans exists. Human plasma N-glycome is quite stable within the individual [6], but is extremely sensitive to pathophysiological processes, reflecting the cell state from the time of protein secretion [7]. The main diagnostic deficit of glycans is in their low specificity, thus they are more frequently studied as prognostic and stratification markers. Their great variability within population, and the significant heritability rate, gives them a great potential in risk assessments. Changes in glycosylation of plasma proteins have been studied in many different diseases, including type 2 diabetes mellitus [8]. These changes have been confirmed in both patients with diabetes and *db/db* mice [9, 10]. Although many studies identified biantennary glycans with core fucose as structures of main interest, differences between cases and healthy population were never pronounced enough to have a diagnostic potential, probably as a result of relatively good glycaemic control in patients, administration of medicaments or existence of many different molecular mechanisms that lead to type 2 diabetes development.

The aim of this study was to determine whether plasma protein N-glycome is changed in individuals who are at greater risk of developing type 2 diabetes, using three different populations:

1. Two populations including individuals with registered hyperglycaemia in critical illness (increased risk for development of type 2 diabetes) and individuals with normoglycaemia in critical illness - AcuteInflammation and AcuteInflammation Replication populations;
2. Individuals who developed type 2 diabetes during 10 years follow-up and those who remained normoglycaemic - FinRisk population.

Also, to further confirm the connection of changes in N-glycome with glucose metabolism disruption, we used data from our previous plasma protein glycosylation analysis in two other

cohorts (ORCADES and SABRE), where we separated age- and sex-matched individuals into two subgroups differing in HbA_{1c} status (HbA_{1c} < 6.5 % (47.5 mmol/mol) and HbA_{1c} > 6.5 % (47.5 mmol/mol)).

Study populations and methods

Patients with/without hyperglycaemia during critical illness (AcuteInflammation and AcuteInflammation Replication populations)

AcuteInflammation population - We included patients from a medical intensive care unit (ICU) at the University Hospital Centre Zagreb during a period of 6 months (February to July 2013). Adult (>18 years old) patients with negative history of diabetes who were admitted to the ICU and discharged from the hospital alive were eligible for inclusion. Excluded were patients diagnosed with diabetes or impaired glucose tolerance and/or impaired fasting glucose before or during hospitalization, patients with documented gestational diabetes, pregnant patients and patients taking glucocorticoids during or 3 months before the admission. Informed consent process was conducted by a member of the study team at the discharge from the ICU or the hospital. Consenting patients were asked to attend a follow-up appointment, 6-8 weeks after the hospital discharge. At this visit inclusion/exclusion criteria were confirmed. Complete blood count and CRP were determined to exclude persisting inflammatory process. Patients with elevated markers of inflammation were retested after 2 weeks. All patients underwent oral glucose tolerance test (OGTT) and measurement of HbA_{1c} to identify pre-existing diabetes. ADA criteria for diagnosis of diabetes were employed and patients diagnosed with existing pre-diabetes or diabetes were excluded. Height, weight and BMI were recorded; family history of diabetes was documented. For all patients fasting blood samples for N-glycan profiling were collected in tubes with anticoagulants (both EDTA and citrate for each patient), plasma was separated immediately and stored at -20°C until the analysis. For each patient, samples were analysed in duplicates. In total, 108 patients were enrolled in the study.

AcuteInflammation Replication population - All cardiac surgical patients operated at the University Hospital Centre Zagreb (cardiac surgical procedures employing cardiopulmonary bypass) from October 2010 to February 2011 and discharged from the hospital alive were screened for the study. All individuals provided written informed consent. Excluded were patients diagnosed with diabetes or impaired glucose tolerance and/or impaired fasting glucose before or during hospitalization. In total, 66 patients were enrolled in the study and from all of them blood was drawn on the day of the surgery, before the surgical procedure. The blood samples were collected in tubes containing EDTA, plasma was separated and stored at -20°C until analysis.

Relevant demographic data for both populations are summarized in Table 2.

Individuals who did/did not develop T2D within 10 years from sampling (FinRisk study)

Plasma samples, stored at - 70°C, collected through the study designed to investigate risk factors in Finland population [11] were used. Population data were searched for patients with incident type 2 diabetes after 10 years follow-up. Individuals who developed any autoimmune, malignant or other chronic disease were excluded. Thirty-seven individuals (aged 60 (34, 72)), who developed type 2 diabetes and no other chronic disease, were identified and included in this study. Thirty-seven age and sex matching controls (age 61 (35, 73)), who remained healthy during the follow-up period, were selected from the same population.

Individuals with different HbA_{1c} status (ORCADES and SABRE studies)

Plasma samples from the Orkney Complex Disease Study (ORCADES), collected in Scottish archipelago of Orkney, between 2005 and 2011 [12] as well as fromSABRE population

were used [13]. Data on age, sex and HbA_{1c} levels were included in the study. Glycan data for aforementioned population had been obtained for another purpose, but was used in this study to compare N-glycomes of individuals with HbA_{1c} < 6.5 % (47.5 mmol/mol) and HbA_{1c} > 6.5 % (47.5 mmol/mol). From Orkney population glycan and HbA_{1c} data for 752 individuals aged 18 to 100 years from a subgroup of ten islands were used for this analysis. Consecutively, 658 individuals (aged 63 (18, 97)) were identified as having HbA_{1c} < 6.5 % (47.5 mmol/mol) and 94 (aged 63 (19, 88)) as having HbA_{1c} > 6.5 % (47.5 mmol/mol). From SABRE population, data for 307 individuals (aged 70 (60, 84)) with HbA_{1c} < 6.5 % (47.5 mmol/mol) and 307 (aged 70 (60, 84)) with HbA_{1c} > 6.5 % (47.5 mmol/mol) were included.

All studies are supported by written informed consent from all individuals and approvals from eligible local Ethics Committees.

Glycan analysis

Glycan release and labelling - Prior to analysis, all samples (cases and controls) have been randomized throughout the multi-well plates. For each plate 4 standards (pool of plasma samples) have been added to minimize the experimental error. Each plasma sample (10 μL) was denatured with the addition of 20 μL of 2 % SDS (w/v) (Invitrogen, Carlsbad, CA, USA) and by incubation at 65 °C for 10 min. Subsequently, 10 μL of 4 % Igepal-CA630 (Sigma-Aldrich, St. Louis, MO, USA) and 1.25 mU PNGase F (ProZyme, Hayward, CA, USA) in 10 μL 5× PBS were added. The samples were incubated overnight at 37 °C for N-glycan release. The released N-glycans were labelled with 2-aminobenzamide (2-AB). The labelling mixture was freshly prepared by dissolving 2-AB (19.2 mg/mL, Sigma-Aldrich, St. Louis, MO, USA) and 2-picoline borane (44.8 mg/mL, Sigma-Aldrich, St. Louis, MO, USA) in DMSO (Sigma-Aldrich, St. Louis,

MO, USA) and glacial acetic acid (Merck, Darmstadt, Germany) mixture (70:30, v/v). Labelling mixture (25 μL) was added to each N-glycan sample in the 96-well plate, which was then sealed using adhesive seal. Mixing was achieved by shaking for 10 min, followed by two-hour incubation at 65 °C. Samples (in a volume of 75 μL) were brought to 80% ACN (v/v) by adding 300 μL of ACN (J.T. Baker, Phillipsburg, NJ, USA). Free label and reducing agent were removed from the samples using HILIC-SPE. To each well of a 0.45 μm GHP filter plate (Pall Corporation, Ann Arbor, MI, USA) 200 μL of 0.1 g/mL suspension of microcrystalline cellulose (Merck, Darmstadt, Germany) in water was added. Solvent was removed by application of vacuum using a vacuum manifold (Millipore Corporation, Billerica, MA, USA). All wells were prewashed using 5× 200 μL of water, followed by equilibration using 3× 200 μL of acetonitrile/water (80:20, v/v). The samples were loaded to the wells, which were subsequently washed 7× using 200 μL of acetonitrile/water (80:20, v/v). Glycans were eluted 2× with 100 μL of water and combined eluates were stored at −20 °C until usage.

Hydrophilic Interaction Chromatography (HILIC)-UPLC - Fluorescently labelled N-glycans were separated by hydrophilic interaction chromatography on Waters Acquity ultra-performance liquid chromatography (UPLC) instrument (Milford, MA, USA) consisting of a quaternary solvent manager, sample manager and a fluorescence detector set with excitation and emission wavelengths of 250 and 428 nm, respectively. The instrument was under the control of Empower 2 software, build 2145 (Waters, Milford, MA, USA). Labelled N-glycans were separated on a Waters bridged ethylene hybrid (BEH) Glycan chromatography column, 150 × 2.1 mm, 1.7 μm BEH particles, with 100 mM ammonium formate, pH 4.4, as solvent A and acetonitrile as solvent B. The separation method used a linear gradient of 70–53% acetonitrile (v/v) at flow rate of 0.561 ml/min in a 25 min analytical run. Samples were maintained at 5 °C

before injection, and the separation temperature was 25 °C. The system was calibrated using an external standard of hydrolysed and 2-AB labelled glucose oligomers from which the retention times for the individual glycans were converted to glucose units. Data processing was performed using an automatic processing method with a traditional integration algorithm after which each chromatogram was manually corrected to maintain the same intervals of integration for all the samples. The chromatograms were all separated in the same manner into 46 peaks (GP1-GP46) and the amount of glycans in each peak was expressed as % of total integrated area (Fig 1).

Statistical analysis

In order to remove experimental variation from measurements, normalisation and batch correction were performed on UPLC glycan data. Also, normalisation by total area was performed in which the peak area of each glycan peak was divided by the total area of the corresponding chromatogram. Prior to batch correction, normalized glycan measurements were log-transformed due to right skewness of their distributions and the multiplicative nature of batch effects. Batch correction was performed on log-transformed measurements using the ComBat method, in which the technical source of variation (which sample was analysed on which plate) was modelled as a batch covariate. To get measurements corrected for experimental noise, estimated batch effects were subtracted from log-transformed measurements.

From the 46 directly measured glycan traits, 12 derived traits were calculated (Table 1).

Table 1. Twelve derived glycan traits calculated from 46 directly measured glycan traits.

Structural feature	Formula
low branching	LB = GP1+GP2+GP3+GP4+GP5+GP6+GP7+GP8+GP9+GP10+GP11+GP12+GP13+GP14+GP15+GP16+GP17+ GP18+GP19+GP20+GP21+GP22+GP23+GP24+GP25+GP26+GP27+GP28
high branching	HB = GP29+GP30+GP31+GP32+GP33+GP34+GP35+GP36+GP37+GP38+GP39+GP40+GP41+GP42+GP43+GP44+GP45+GP46
agalactosylated	G0 = GP1+GP2+GP4+GP5+GP6
monogalactosylated	G1 = GP3+GP7+GP8+GP9+GP10+GP16+GP17+GP18
digalactosylated	G2 = GP12+GP13+GP14+GP15+GP19+GP20+GP21+GP22+GP23+GP25+GP26+GP27+GP28
trigalactosylated	G3 = GP29+GP31+GP32+GP33+GP34+GP35+GP36+GP37+GP40
tetragalactosylated	G4 = GP30+GP38+GP39+GP41+GP42+GP43+GP44+GP45+GP46
neutral	S0 = GP1+GP2+GP3+GP4+GP5+GP6+GP7+GP8+GP9+GP10+GP11+GP12+GP13+GP14+GP15
monosialylated	S1 = GP16+GP17+GP18+GP19+GP20+GP21+GP22+GP30
disialylated	S2 = GP23+GP25+GP26+GP27+GP28+GP29+GP31
trisialylated	S3 = GP32+GP33+GP34+GP35+GP36+GP37+GP38+GP39+GP40
tetrasialylated	S4 = GP41+GP42+GP43+GP44+GP45+GP46

These derived traits average particular glycosylation features (branching, galactosylation, sialylation) across different individual glycan structures, and consequently they are more closely related to individual enzymatic activities and underlying genetic polymorphisms. As derived traits represent sums of directly measured glycans, they were calculated using normalized and batch-corrected glycan measurements after transformation to the proportions (exponential transformation of batch-corrected measurements).

Analyses of associations between clinical trait of interest (AcuteInflammation and AcuteInflammation Replication - hyperglycaemia; FinRisk Cohort - future onset of diabetes; ORCADES & SABRE Cohorts - HbA_{1C} > 6.5 % (47.5 mmol/mol)) and glycan measurements were performed using a regression model with age and sex included as additional covariates. Prior to analyses, for each cohort separately, glycan variables were transformed to a standard normal distribution by inverse transformation of ranks to normality. Using rank-transformed

variables makes estimated effects of different glycans in different cohorts comparable, as transformed glycan variables have the same standardized variance. The false discovery rate (FDR) was controlled using the Benjamini-Hochberg procedure.

Data were analysed and visualized using R programming language (version 3.0.1).

Results

Analysis of plasma N-glycome composition in AcuteInflammation and AcuteInflammation Replication populations

For each patient, N-glycans were released from plasma proteins, fluorescently labelled and analysed by HILIC-UPLC, a method that was demonstrated to be the best approach for reliable and reproducible quantitative glycan analysis [14]. This method separates plasma glycome into 46 chromatographic peaks, each containing a group of similar glycan structures. Individual glycan structures contained in each peak were characterized previously [15] and representative chromatogram with main structure present in each peak is shown in Fig 1.

Next, we quantified each chromatographic peak (as % of the total N-glycome) and calculated derived traits (Table 1). These derived traits represent common biologically meaningful features (glycan branching, galactosylation and sialylation) shared among several measured glycans. Then we compared these traits between two groups of patients in two populations:

1. Those who developed hyperglycaemia during ICU hospitalization due to acute condition and those who remained normoglycaemic during the same circumstances (AcuteInflammation population, Table 2) – for the analysis, plasma samples taken 6-8 weeks after the hospital discharge were used, and meanwhile the absence of inflammatory process was also confirmed (acquired profiles thus represented basal N-glycomes for each person);
2. Those who developed hyperglycaemia after cardiac surgery due to acute inflammation and those who remained normoglycaemic after the surgery (AcuteInflammation Replication population, Table 2) – for the analysis, plasma samples taken on the day of the surgery, before the surgical procedure were used.

The development of hyperglycaemia during acute inflammation represents significantly higher risk for type 2 diabetes development [2, 3].

Fig 2 and Table 3 show differences in abundance of derived structural features of plasma protein N-glycans in individuals who developed hyperglycaemia during acute condition and controls. Differences are shown as the effect sizes estimated based on logistic regression (logarithm of odds ratio). In AcuteInflammation population, six derived traits were significantly different between the two groups – patients who developed hyperglycaemia had decreased low branching (LB) and increased high branching (HB), increased tri- and tetragalactosylation (G3 and G4), decreased neutral glycans (S0) and increased trisialylation (S3).

The differences found in AcuteInflammation Replication population showed the same trend and similar magnitude of changes, as found in the AcuteInflammation population for almost all derived traits, except for digalactosylation (G2). However, none of them reached statistical significance.

Analysis of plasma N-glycome composition in FinRisk populations

Differences in plasma protein N-glycome between individuals who will or will not develop type 2 diabetes was further valuated in the FinRisk study cohort, in which plasma samples were collected more than 10 years ago and biobanked. The 10-year follow-up has been performed to note the incidence of type 2 diabetes. Thirty-seven individuals who developed type 2 diabetes and no other chronic disease and 37 age and sex matching controls, who remained healthy during follow-up period, were identified and included in the study. Plasma N-glycome was analysed in the same way as for the previous two populations.

Individuals who developed type 2 diabetes had significantly decreased LB and increased HB, decreased agalactosylation and monogalactosylation (G0 and G1), decreased neutral glycans (S0) and increased di- and tri-sialylation (S2 and S3), compared to individuals who remained normoglycaemic (Fig 2 and Table 3). For all derived glycan traits, differences found in FinRisk were in the same direction and of similar magnitude, as found in AcuteInflammation population.

Analysis of plasma N-glycome composition in ORCADES and SABRE populations

To further prove the connection of N-glycome with glucose metabolism disruption, we used data from our previous analysis of plasma protein glycosylation in two other cohorts (ORCADES and SABRE), where we separated age- and sex-matched individuals into two subgroups differing in HbA_{1c} status (HbA_{1c} < 6.5 % (47.5 mmol/mol) and HbA_{1c} > 6.5 % (47.5 mmol/mol)).

We compared the differences between the two subgroups in the same way as in the previous populations. Levels of the majority of glycan features were significantly different between the studied subgroups (10 for ORCADES and 7 for SABRE) (Fig 2 and Table 3). Main differences were the same as in the previous populations - higher levels of branching, galactosylation and sialylation in their plasma protein N-glycomes.

Conclusions

Our results indicate that plasma protein N-glycome is changed in individuals with increased risk for type 2 diabetes development (AcuteInflammation and AcuteInflammation Replication populations), incident cases of type 2 diabetes collected at baseline (FinRisk population) and people with increased HbA_{1c} (ORCADES and SABRE populations). Differences in plasma N-glycomes were the same in all three populations and included higher levels of

branching, galactosylation and sialylation. This suggests that the increased complexity of glycan structures represents the greater chance for diabetes type 2 development and is also associated with poorer regulation of blood glucose levels.

One can only speculate whether these differences in plasma glycome are inborn and genetically conditioned, or glycosylation is affected by the pathophysiological mechanisms that occur very early at diabetes onset and manifest themselves in the changes of the plasma glycome composition due to high susceptibility of glycans to changes in the cell metabolism. A possible explanation for this higher branching could be that higher glycaemia leads to altered flux through the hexosamine pathway, which produces UDP-GlcNAc, the substrate for N-linked glycosylation [16]. Recently, it has also been shown that the hexosamine biosynthesis and GlcNAc salvage pathways contribute to glucose homeostasis through N-glycan branching on glucagon receptor [17].

Previous studies showed that multi-branched and highly sialylated N-glycans were also elevated in response to inflammatory diseases, such as ulcerative colitis [18], chronic pancreatitis [19] and rheumatoid arthritis [20]. It is also well known that individuals with the metabolic syndrome and type 2 diabetes suffer from chronic low-grade inflammation [21]. Therefore, the N-glycan changes that were observed in this study may also reflect the chronic inflammatory processes. Additional support for this hypothesis comes from our recent study of a NMR biomarker GlycA and total plasma protein N-glycome. GlycA is a signal that measures N-acetylglucosamines attached to the plasma proteins and it has been shown to correlate with wide spectra of inflammatory diseases, including incident type 2 diabetes mellitus [22, 23]. In that study we found that GlycA signal was positively related to highly branched structures, as well as tri- and tetragalactosylated and tri- and tetrasialylated structures, while low branched glycans,

agalactosylated, monogalactosylated, asialylated and monosialylated structures were negatively associated with GlycA signal [24]. These changes are practically the same as found herein, which suggests that the changes in plasma protein N-glycome could indicate the start of chronic inflammation and the susceptibility for developing the metabolic syndrome. This could also explain why the differences in derived glycan traits between the cases and controls in AcuteInflammation Replication population did not reach the threshold of significance – all individuals from this population had cardiovascular diseases and underwent cardiac surgery. Therefore, it is possible that most of them (both cases and controls) already had some kind of chronic inflammation. The other possible reason for this could be that the control group was too small (14 individuals).

Despite the fact that we measured the whole plasma protein N-glycome, which is comprised of different glycans originating from many different glycoproteins, we managed to find differences in glycan traits between the controls and seemingly healthy individuals, who have a higher chance or will develop type 2 diabetes mellitus. The same changes were also associated with poorer regulation of blood glucose. Identifying the exact glycoproteins which contribute to these differences would probably help to develop stratification methods that could reliably distinguish individuals who are at risk of type 2 diabetes development and improve the prevention of this widespread disease and its complications.

Data availability: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request**.**

Funding: This study was supported by FP7 grants IntegraLife (contract #315997 to GL) and HighGlycan (contract # 278535 to GL), as well as The University of Zagreb supporting grant (to OG) and Australian-China Collaborative Grant (NH&MRC- APP1112767 -NSFC 81561128020 to WW). The SABRE study was funded at baseline by the UK Medical Research Council, Diabetes UK, and the British Heart Foundation, and at follow-up by the Wellcome Trust (WT082464 to NC), British Heart Foundation (SP/07/001/23603 and CS/13/1/30327 to NC), Diabetes UK(13/0004774 to NC). The FinRisk study was mainly funded by budgetary funds from the National Institute for Health and Welfare. Important additional funding has been obtained from the Academy of Finland (grant # 139635 to VS) and from the Finnish Foundation for Cardiovascular Research.

Duality of interest: The authors declare that there is no duality of interest associated with this manuscript.

Contribution statement: OG, IGo, and GL conceived and designed the study; OG, GL, HC and WW supervised the study; TK, FV, NS, IGu, TP, EL, HG, BB, TT, AW, SM, VS, AH, JFW, NC, MP participated in data acquisition, collection, analysis or interpretation; OG, TK, and FV drafted the manuscript; GL, NS, IGo, IGu, TP, EL, HG, BB, TT, AW, SM, VS, AH, JFW, NC, MP, HC and WW critically revised the manuscript for intellectual content. All authors approved the final version of the manuscript. OG is the guarantor of the study and is responsible for the integrity of the work as a whole.

Figure legends

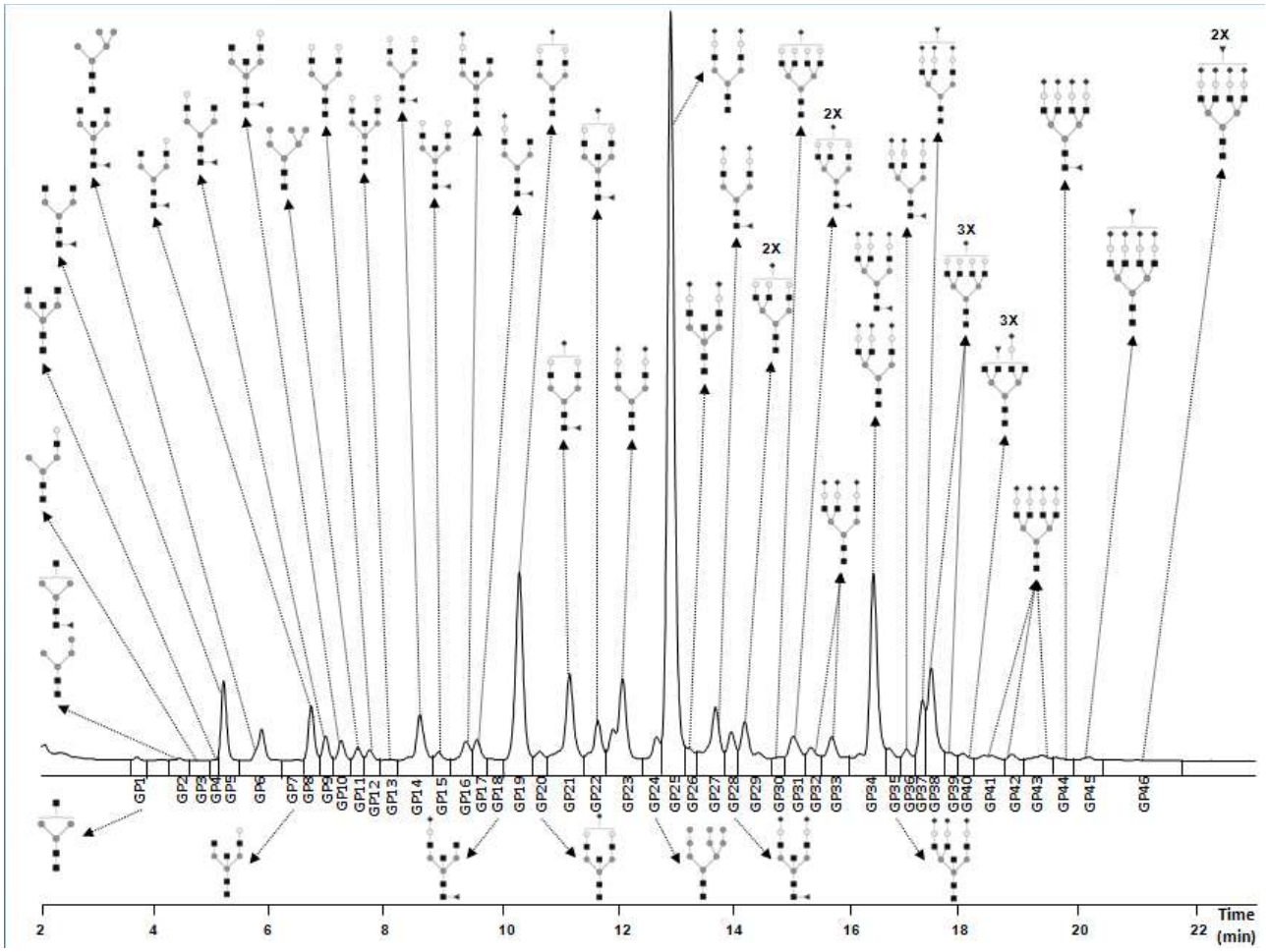

Figure 1. HILIC-UPLC profile of N-linked glycans released from plasma proteins and fluorescently labelled with 2-aminobezamide. Glycans presented are the most abundant structures in each peak. Structural schemes: square, N-acetylglucosamine; triangle, fucose; dark grey circle, mannose; light grey circle, galactose; diamond, N-acetylneuraminic acid.

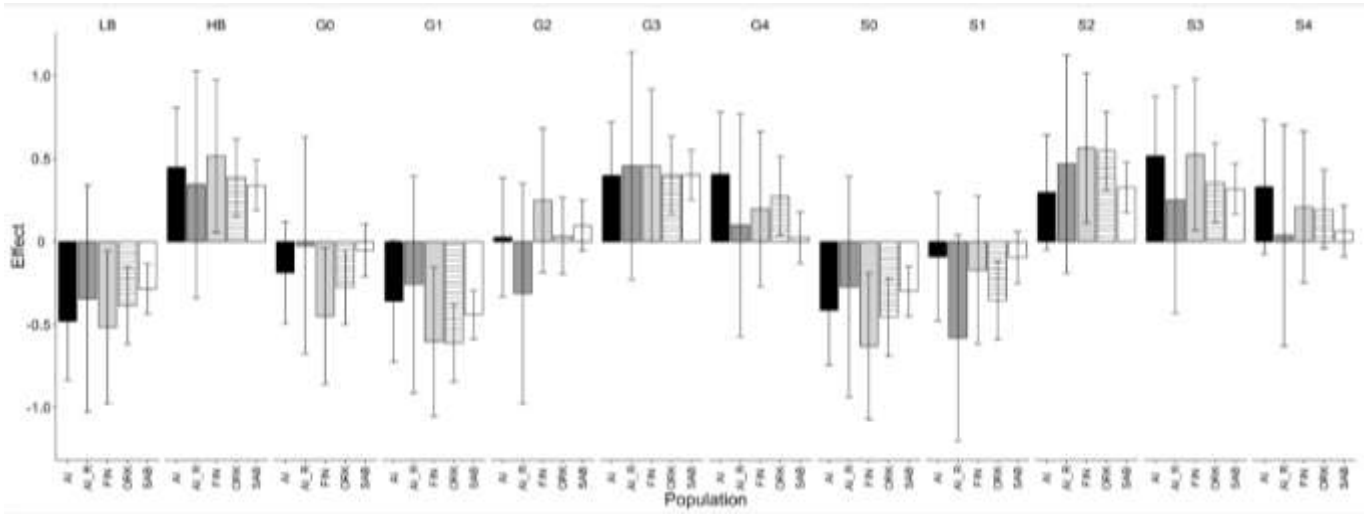

Figure 2. Differences in abundance of structural features of plasma protein N-glycans in individuals who developed hyperglycaemia during acute condition (AcuteInflammation and AcuteInflammation Replication populations - black and dark grey bars, respectively); in individuals from FinRisk population who developed type 2 diabetes during 10-year follow-up (light grey bars); in individuals from ORKNEY population (bars with horizontal stripes) and SABRE population (white bars) divided according to HbA_{1c} level (< or > 6.5 % (47.5 mmol/mol)). Pronounced differences were observed between cases and controls in all cohorts for several derived glycan structural features. Differences in derived glycan traits are shown as bar plots where height of bars represents the size of effects estimated based on logistic regression (logarithm of odds ratio) and error bars represent 95 % confidence intervals of estimated effects. (LB - low branching, sum of monoantennary and biantennary glycans, see Methods section; HB - high branching, sum of triantennary and tetraantennary glycans; G0 - sum of glycans with no galactose; G1 - sum of glycans with one galactose; G2 - sum of glycans with two galactoses; G3

- sum of glycans with three galactoses; G4 - sum of glycans with four galactoses; S0 - sum of glycans with no sialic acid; S1 - sum of glycans with one sialic acid; S2 - sum of glycans with two sialic acids; S3 - sum of glycans with three sialic acids; S4 - sum of glycans with four sialic acids; AI - AcuteInflammation; AI_R - AcuteInflammation Replication; FIN - FinRisk; ORK - ORKNEY; SAB - SABRE).

References

1. International Diabetes Federation (IDF). IDF diabetes Atlas 7th edition. idf.org. 2015: available from: http://www.diabetesatlas.org accessed 15 June 2016
2. Gornik I, Vujaklija A, Lukic E, Madzarac G, Gasparovic V: Hyperglycemia in sepsis is a risk factor for development of type II diabetes. J Crit Care 2010;25:263-269
3. Gornik I, Vujaklija A, Lukic E, Madzarac G, Gasparovic V: Hyperglycaemia in critical illness is a risk factor for later development of type II diabetes mellitus. Acta Diabetol 2010;47 Suppl 1:29-33
4. Gornik O, Pavic T, Lauc G: Alternative glycosylation modulates function of IgG and other proteins - implications on evolution and disease. Biochim Biophys Acta 2012;1820:1318-1326
5. Lauc G, Rudan I, Campbell H, Rudd PM: Complex genetic regulation of protein glycosylation. Mol Biosyst 2010;6:329-335
6. Gornik O, Wagner J, Pucic M, Knezevic A, Redzic I, Lauc G: Stability of N-glycan profiles in human plasma. Glycobiology 2009;19:1547-1553
7. Drake PM, Cho W, Li B, Prakobphol A, Johansen E, Anderson NL, Regnier FE, Gibson BW, Fisher SJ: Sweetening the pot: adding glycosylation to the biomarker discovery equation. Clin Chem 2010;56:223-236
8. Testa R, Vanhooren V, Bonfigli AR, Boemi M, Olivieri F, Ceriello A, Genovese S, Spazzafumo L, Borelli V, Bacalini MG, Salvioli S, Garagnani P, Dewaele S, Libert C, Franceschi C: N-glycomic changes in serum proteins in type 2 diabetes mellitus correlate with complications and with metabolic syndrome parameters. PLoS One 2015;10:e0119983
9. Itoh N, Sakaue S, Nakagawa H, Kurogochi M, Ohira H, Deguchi K, Nishimura S, Nishimura M: Analysis of N-glycan in serum glycoproteins from db/db mice and humans with type 2 diabetes. Am J Physiol Endocrinol Metab 2007;293:E1069-1077
10. Higai K, Azuma Y, Aoki Y, Matsumoto K: Altered glycosylation of alpha1-acid glycoprotein in patients with inflammation and diabetes mellitus. Clin Chim Acta 2003;329:117-125
11. Vartiainen E, Laatikainen T, Peltonen M, Juolevi A, Mannisto S, Sundvall J, Jousilahti P, Salomaa V, Valsta L, Puska P: Thirty-five-year trends in cardiovascular risk factors in Finland. Int J Epidemiol 2010;39:504-518
12. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A, Macleod AK, Farrington SM, Rudan P, Hayward C, Vitart V, Rudan I, Wild SH, Dunlop MG, Wright AF, Campbell H, Wilson JF: Runs of homozygosity in European populations. Am J Hum Genet 2008;83:359-372
13. Tillin T, Forouhi NG, McKeigue PM, Chaturvedi N: Southall And Brent REvisited: Cohort profile of SABRE, a UK population-based comparison of cardiovascular disease and diabetes in people of European, Indian Asian and African Caribbean origins. Int J Epidemiol 2012;41:33-42
14. Huffman JE, Pucic-Bakovic M, Klaric L, Hennig R, Selman MH, Vuckovic F, Novokmet M, Kristic J, Borowiak M, Muth T, Polasek O, Razdorov G, Gornik O, Plomp R, Theodoratou E, Wright AF, Rudan I, Hayward C, Campbell H, Deelder AM, Reichl U, Aulchenko YS, Rapp E, Wuhrer M, Lauc G: Comparative performance of four methods for high-throughput glycosylation analysis of immunoglobulin G in genetic and epidemiological research. Mol Cell Proteomics 2014;13:1598-1610

15. Saldova R, Asadi Shehni A, Haakensen VD, Steinfeld I, Hilliard M, Kifer I, Helland A, Yakhini Z, Borresen-Dale AL, Rudd PM: Association of N-glycosylation with breast carcinoma and systemic features using high-resolution quantitative UPLC. J Proteome Res 2014;13:2314-2327
16. Ohtsubo K, Chen MZ, Olefsky JM, Marth JD: Pathway to diabetes through attenuation of pancreatic beta cell glycosylation and glucose transport. Nat Med 2011;17:1067-1075
17. Johswich A, Longuet C, Pawling J, Abdel Rahman A, Ryczko M, Drucker DJ, Dennis JW: N-glycan remodeling on glucagon receptor is an effector of nutrient sensing by the hexosamine biosynthesis pathway. J Biol Chem 2014;289:15927-15941
18. Miyahara K, Nouso K, Saito S, Hiraoka S, Harada K, Takahashi S, Morimoto Y, Kobayashi S, Ikeda F, Miyake Y, Shiraha H, Takaki A, Okada H, Amano M, Hirose K, Nishimura S, Yamamoto K: Serum glycan markers for evaluation of disease activity and prediction of clinical course in patients with ulcerative colitis. PloS one. 2013;8(10):e74861
19. Sarrats A, Saldova R, Pla E, Fort E, Harvey DJ, Struwe WB, de Llorens R, Rudd PM, Peracaula R: Glycosylation of liver acute-phase proteins in pancreatic cancer and chronic pancreatitis. Proteom Clin Appl. 2010;4(4):432-48
20. Pawlowski T, Mackiewicz A, Mackiewicz S: Studies on microheterogeneity of acute-phase proteins in rheumatoid arthritis by using crossed affinoimmuno-electrophoresis with free concanavalin A. Behring Inst Mitt. 1986;(80):11-5
21. Eckel RH, Grundy SM, Zimmet PZ: The metabolic syndrome. Lancet. 2005;365(9468):1415-28
22. Otvos JD, Shalaurova I, Wolak-Dinsmore J, Connelly MA, Mackey RH, Stein JH, Tracy RP: GlycA: A Composite Nuclear Magnetic Resonance Biomarker of Systemic Inflammation. Clin Chem 2015;61:714-723
23. Akinkuolie AO, Pradhan AD, Buring JE, Ridker PM, Mora S: Novel protein glycan side-chain biomarker and risk of incident type 2 diabetes mellitus. Arterioscler Thromb Vasc Biol 2015;35:1544-1550
24. Gudelj I, Vučković F, Tilin T, Mora S, Keser T, Akinkuolie AO, Trbojević-Akmačić I, Pučić-Baković M, Chaturvedi N, Lauc G: Total plasma N-glycome profiling is a more predictive biomarker than GlycA NMR signal. In the process of publishing

Tables

Table 2. The demographic characteristics of AcuteInflammation and AcuteInflammation Replication populations.

	All patients		HG during acute illness		No HG during acute illness		P value (HG vs. no HG)[a]	
Population	AI	AI_R	AI	AI_R	AI	AI_R	AI	AI_R
Number of patients (N)	108	66	59	52	49	14		
Age (y)	55 (18, 79)	67 (21, 81)	58 (25, 79)	68 (41, 81)	54 (18, 75)	62 (21, 78)	0.388	0.121
Female sex (N,%)	32 (29.6%)	19 (28.8%)	18 (30.5%)	15 (28.8%)	14 (28.6%)	4 (28.6%)	0.826	0.984
BMI (kg/m2)	23.7±4.0	28.2±3.5	24.2±4.1	28.5±3.4	23.1±3.9	27.2±3.8	0.734	0.342
Family history of diabetes (N,%)	27 (25.0%)	NA	19 (32.2%)	NA	8 (16.3%)	NA	0.039	NA

(AI - AcuteInflammation; AI_R - AcuteInflammation Replication; HG - hyperglycaemia; BMI - body mass index; [a] P values were calculated using Mann-Whitney U test (significance level α=0.05) and for categorical variables Chi-square test was used)

Table 3. Differences in derived structural features of plasma protein N-glycans between the cases and the controls in all five populations. Differences are shown as effects estimated based on logistic regression (logarithm of odds ratio).

	AcuteInflammation				AcuteInflammation_R				FinRisk				ORCADES				SABRE			
derived glycan trait	effect	SE	P value	adj. P value	effect	SE	P value	adj. P value	effect	SE	P value	adj. P value	effect	SE	P value	adj. P value	effect	SE	P value	adj. P value
LB	-0.480	0.179	0.006	0.016*	-0.344	0.342	0.281	0.393	-0.517	0.230	0.022	0.040*	-0.384	0.119	0.001	0.002**	-0.284	0.077	0.000	0.001**
HB	0.448	0.180	0.010	0.024*	0.344	0.342	0.281	0.393	0.517	0.230	0.022	0.040*	0.388	0.119	0.001	0.002**	0.339	0.077	0.000	0.000***
G0	-0.186	0.154	0.209	0.276	-0.023	0.327	0.940	0.956	-0.447	0.206	0.027	0.045*	-0.276	0.113	0.014	0.019*	-0.052	0.080	0.511	0.558
G1	-0.359	0.184	0.044	0.068	-0.257	0.326	0.399	0.498	-0.603	0.225	0.007	0.017*	-0.611	0.119	0.000	0.000***	-0.440	0.074	0.000	0.000***
G2	0.027	0.182	0.878	0.878	-0.312	0.332	0.314	0.428	0.250	0.217	0.236	0.290	0.036	0.118	0.759	0.759	0.097	0.079	0.213	0.276
G3	0.399	0.163	0.011	0.024*	0.458	0.343	0.153	0.236	0.459	0.230	0.041	0.066	0.403	0.118	0.001	0.002**	0.403	0.076	0.000	0.000***
G4	0.406	0.190	0.027	0.045*	0.099	0.335	0.750	0.818	0.198	0.234	0.381	0.446	0.276	0.120	0.021	0.025*	0.024	0.079	0.756	0.789
S0	-0.413	0.167	0.011	0.024*	-0.271	0.332	0.380	0.487	-0.630	0.221	0.004	0.012*	-0.455	0.119	0.000	0.001***	-0.297	0.077	0.000	0.001***
S1	-0.091	0.196	0.627	0.669	-0.581	0.310	0.047	0.083	-0.171	0.224	0.429	0.478	-0.354	0.120	0.003	0.005**	-0.096	0.080	0.230	0.290
S2	0.298	0.174	0.077	0.111	0.468	0.329	0.129	0.204	0.565	0.226	0.011	0.024*	0.547	0.120	0.000	0.000***	0.327	0.077	0.000	0.000***
S3	0.517	0.181	0.003	0.010*	0.252	0.341	0.428	0.514	0.525	0.230	0.020	0.039*	0.358	0.121	0.003	0.005**	0.319	0.077	0.000	0.000***
S4	0.331	0.204	0.093	0.128	0.038	0.333	0.902	0.933	0.209	0.229	0.347	0.416	0.195	0.122	0.106	0.115	0.063	0.079	0.424	0.478

(LB - low branching, sum of monoantennary and biantennary glycans, see Methods section; HB - high branching, sum of triantennary and tetraantennary glycans; G0 - sum of glycans with no galactose; G1 - sum of glycans with one galactose; G2 - sum of glycans with two galactoses; G3 - sum of glycans with three galactoses; G4 - sum of glycans with four galactoses; S0 - sum of glycans with no sialic acid; S1 - sum of glycans with one sialic acid; S2 - sum of glycans with two sialic acids; S3 - sum of glycans with three sialic acids; S4 - sum of glycans with four sialic acids; SE - standard error; adj. P value – P value corrected for multiple measures using Benjamini-Hochberg procedure; *P < 0.05, **P < 0.01, ***P < 0.001)