Global glycoproteomic analysis reveals specific glycation and glycosylation events associated with cognitive decline in older adults with type 2 diabetes

Nili Tickotsky (nilitiko@gmail.com)  
Bar Ilan University

David Morgenstern  
Nancy and Stephen Grand Israel National Center for Personalized Medicine, Weizmann Institute of Science

Hila Wolf-Levy  
Nancy and Stephen Grand Israel National Center for Personalized Medicine, Weizmann Institute of Science

Itzik Cooper  
Sheba Medical Center

Jaime Uribarri  
The Icahn School of Medicine at Mount Sinai

Ron Unger  
Bar Ilan University

Aron S. Buchman  
Rush University Medical Center

Yishai Levin  
Nancy and Stephen Grand Israel National Center for Personalized Medicine, Weizmann Institute of Science

Michal Schnaider Beeri  
The Icahn School of Medicine at Mount Sinai

Article

Keywords: Cognitive impairment, Glycation, Glycosylation, Mass spectrometry, Type 2 diabetes

Posted Date: November 4th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2205363/v1
Abstract

INTRODUCTION: Molecular mechanisms underlying the well-known association between Type 2 diabetes (T2D) and late-life cognitive impairment are unknown. Alterations in glyco post-translational modifications (PTM) of serum proteins reflect altered cellular physiology and are associated with T2D and impaired cognition.

This novel large-scale mass spectrometry based glycoproteomics study aimed to identify therapeutic PTM targets that may link T2D with poor cognition.

METHODS: We performed LC-MS/MS–based novel glycoproteomic profiling of sera samples from 23 cognitively normal older adults with T2D. At baseline all people had normal cognition, but during 36-months, 8 participants developed cognitive impairment while 15 maintained normal cognition.

RESULTS: We identified and quantified a total of 1,866 unique glycoproteoforms (i.e., peptides with a post-transcription sugar modification) that matched 201 annotated proteins. Group differences were observed for 72 glycoproteoforms at baseline. At the time of cognitive testing three years later, there were group differences for 137 glycoproteoforms. Group differences were observed for 14 glycoproteoforms both at baseline and at 3-year follow-up.

DISCUSSION: Glycoproteoforms may contribute to cognitive impairment in T2D, may identify adults with T2D at risk for cognitive impairment and are high value therapeutic targets for new therapies to mitigate the negative cognitive effects of T2D in old age.

Introduction

A large body of literature supports the notion that type 2 diabetes may contribute to late-life cognitive impairment\(^1-5\). Yet, the molecular mechanisms underlying this association are unclear\(^5-7\). Leveraging new genomic technologies including transcriptomics and proteomics has led to the identification of novel molecular mechanisms i.e., genes and proteins that may drive late-life cognitive impairment\(^8\). While the protein's amino acid sequence is coded in the genome, posttranslational modifications (PTMs) commonly alter protein functions by changing their physical and chemical properties. While thousands of proteins are interrogated by current proteomic studies, the associations of varied PTM remains unexplored\(^9\).

The most common posttranslational modification of proteins is the addition of carbohydrates \(^9,10\). These modified proteins, (glycoproteoforms) have crucial roles in almost all cellular processes, including cell signaling, immune recognition, and cell-cell interaction \(^11-13\). The structures of the carbohydrates are much more variable than those of proteins and nucleic acids, and they can be altered by changes in the physiological conditions of the organism and the cells \(^12\). Consequently, the heterogeneity of these carbohydrates may contribute to diverse age-related diseases as well as provide a sensitive indicator for age-related cellular changes and diverse diseases \(^9,11,14\).
The carbohydrates posttranslational modifications are created by two types of sugar-adding reactions: Glycation and glycosylation\textsuperscript{15}: Glycosylation is a highly regulated enzymatic reaction that is part of normal protein biosynthesis\textsuperscript{15}, and creates glycan structures. It is the most abundant post-translational protein modification encountered in nature\textsuperscript{9,13}, and the mammalian glycome repertoire contains thousands of glycan structures\textsuperscript{13}. Glycation, on the other hand, is a non-enzymatic, non-regulated reaction associated with diseases or pathological changes related to hyperglycemia in old age\textsuperscript{15–17}.

While animal and human studies suggest that proteins that have undergone glycosylation may contribute to and Alzheimer’s disease and related disorders (ADRD), there is a paucity of data about serum glycoproteomics in and late-life cognitive impairment\textsuperscript{11–13,18}. Technical challenges in measuring glycoproteoforms have impeded studies of the glycoproteome. To address this gap, we published a novel LC-MS/MS–based methodology that can identify proteins modified by both glycation and glycosylation and the locations where glycans link to proteins\textsuperscript{19}. We refer to the specific amino acid sequence i.e., protein together with its linked glycan as a “glycoproteoform”. As a proof of concept, we applied this novel methodology to profile the glycoproteome in serum from older adults with type 2 diabetes to test the hypothesis that serum glycoproteoforms are associated with late-life cognitive impairment in older adults with type 2 diabetes.

**Results**

**Description of the analytic cohort at baseline**

Demographic, clinical and cognitive tests were similar in adults with and without cognitive impairment (Table I). At baseline, the global cognitive scores of those with cognitive impairment (Mean= 0.07, SD=0.56) and those without cognitive impairment (Mean=0.33, SD= 0.44) did not differ significantly (Kolmogorov-Smirnov test, p-value =0.104), but those with cognitive impairment had lower global cognitive scores at 18-months follow ups (those with cognitive impairment mean of global cognition= -0.7 [SD= 0.53] vs. those without cognitive impairment mean= 0.09, [SD= 0.43]) and 36-months follow up (Mean= 0.076 [SD= 0.57] vs. -1.15 [SD= 0.76])(Kolmogorov-Smirnov test, p-value= 0.0006 and 0.001, respectively).

**Baseline glycoproteome results**

We quantified 1,866 unique glycoproteoforms (in at least 30% of the 23 samples and in at least one of the two groups) from 201 proteins (Table S1). In addition, 177 glycoproteoforms had two glycosylation sites (motif based) on a single peptide. In these cases, the individual mass of each of the two glycans was unknown as the mass represents the sum of the two glycans. Therefore, we included these peptides only for statistical analysis but did not use them in the analyses of specific glycan composition types.

While most of the 201 proteins had a single glycoproteoform, several proteins had more than 40 glycoproteoforms (Fig. 1). For example, haptoglobin had 171 glycoproteoforms. Figure 1 (rectangle)
shows two of the haptoglobin glycoproteoforms we found where glycans with different compositions can be attached to the same haptoglobin peptide.

**Baseline glycoproteoforms profiles and cognitive status at follow-up**

Baseline cognitive testing in all adults showed normal cognition (Table 1), but 8 of the 23 had developed cognitive impairment during 3-year follow-up. Group differences were observed in the expression of 43 glycoproteoforms of 37 proteins (Supplementary Table S2). When these 43 glycoproteoforms were grouped based on their biological function using the DAVID protein annotation tool (Fig. 2A), they were found to participate in immune pathways, proteolysis, and blood coagulation. Annotations for the glycated and glycosylated proteins separately are shown in supplementary figures S1, S2.

**Glycoproteoforms profiles and cognitive status at 3-yr follow-up**

Group differences were observed in the relative abundance of 137 glycoproteoforms from 66 proteins at 3-year follow-up (Supplementary Table S3). When grouped by the DAVID protein annotation tool according to their biological function, the main pathways found for these proteins were similar to baseline and included immune system related pathways, blood coagulation and proteolysis (Fig. 2B).

Group differences were observed for 14 glycoproteoforms, nine glycosylated and five glycated, that differed at baseline and in the relative abundance at 3-year follow-up. (Table S4). The main functional pathway of these 14 glycoproteoforms based on annotation with DAVID was the immune system. This included innate and adaptive immune responses, and immunoglobulin receptor binding.

**Immunoglobulin glycosylation**

Five glycosylated glycopeptiforms from five immunoglobulins, IGHG1, IGHG2, IGHG3, IGHG4 and IGHA2, significantly changed between converters and non-converters from baseline to 36 months (Fig. 3).

**Discussion**

This study highlights the utility of a novel glycoproteomics methodology that has potential to facilitate large-scale glycoproteome studies to discover the contribution of post-translational modifications due to glycosylation and glycation to late-life cognitive impairment and other diseases. We profiled sera at baseline and three years later in older adults with type 2 diabetes and normal cognitive function at baseline and compared the results between individuals that maintained normal cognition versus those who developed cognitive impairment. While all adults had normal cognition at baseline we found group differences in glycoproteoforms expression profiles both at baseline and at 3-year follow-up. Serum glycoproteoforms have potential as novel disease biomarkers. Moreover, site-specific glycoproteomic modifications associated with cognitive decline may provide high value therapeutic targets for further drug discovery and studies that seek to understand the mechanisms driving late-life cognitive impairment in type 2 diabetes.
Prior glycomics-based profiling have reported several N-glycan structures that may distinguish and controls\textsuperscript{20,21} and play a role in cognitive impairment pathogenesis and progression\textsuperscript{22}. Recent evidence shows that abnormal central nervous system glycosylation due to alterations in N and O-glycan structures, is implicated in cognitive impairment and neurodegeneration\textsuperscript{11,14,23}. Yet, studies using glycomics do not provide information on the specific protein sites and lack specificity as glycans are summed from all proteins. It has been suggested that analysis of glycans attached to their peptides (i.e., glycoproteomics) may identify glycosylation changes specific to particular proteins and thus provide new insights into the molecular mechanisms underlying the consequences of type 2 diabetes, including cognitive impairment\textsuperscript{21}. One example for such consequence on the immune system is our finding of glycosylated glycopeptiforms from five immunoglobulins that significantly changed between converters and non-converters from baseline to 36 months.

Only a minority of the glycoproteoforms differences between the groups were observed at baseline and at 3-year follow-up. Nonetheless, functional annotation of the glycoproteoforms showed that both at baseline and follow-up proteins that were differentially associated with cognitive status had similar functions including immune system, proteolysis and blood coagulation. Major components of the immune system differentiated between those with cognitive impairment and those without it: the complement pathway, B-cell activation, immunoglobulins (IgG) function and phagocytosis. These immune system components show a discrepancy between the type of protein modifications, with glycated proteins enriched for the complement pathway and glycosylated proteins affecting B-cell activation, immunoglobulins (IgG) function and phagocytosis.

We found 14 glycoproteoforms that differed between the groups at both baseline and follow-up even though all individuals had normal cognition at baseline. Interestingly, their main functional pathways were also linked to innate and adaptive immune responses, and immunoglobulin receptor binding. Each of these components of the immune system has been implicated in cognitive decline and AD\textsuperscript{24–27}. Our study adds new evidence pointing to a role of glycan-related post translational modifications of immune system proteins in cognitive impairment. These glycoproteoforms warrant additional targeted glycoproteomic studies to validate their contribution to old-age type 2 diabetes-related cognitive impairment. They point to the immune system as underlying mechanism and their potential use as biomarkers of individuals at risk of future cognitive impairment.

Ten percent of the proteins showing differentially altered glycoproteoforms belong to the proteolysis pathway. This pathway is damaged by glycative stress\textsuperscript{28,29} and by glycosylation modification, that may contribute to decreased protein degradation\textsuperscript{29}. Chronic hyperglycemia in generates advanced glycation end products (AGEs) and activates their receptor (RAGE)\textsuperscript{30}. RAGE proteolysis regulates RAGE signal transduction and also Aβ peptide transport across the blood-brain-barrier\textsuperscript{31}. Further work is needed to test how damage to the proteolytic pathway by glycan-dependent proteolysis alterations may contribute to diabetes type 2 -related cognitive impairment.
An additional 10% percent of the proteins showing differentially altered glycoproteoforms belong to the blood coagulation pathways. Cerebrovascular disease is the most common pathology underlying cognitive impairment in type 2 diabetes\textsuperscript{32,33}. Blood coagulation pathways have been linked to initiation or aggravation of brain hypoperfusion and brain vascular inflammation\textsuperscript{34}, and there is evidence from AD-like animal models, suggesting that the inhibition of coagulation in AD may serve as a new therapeutic target\textsuperscript{34,35}.

Glycation end product levels are higher in type 2 diabetes, and have been shown to cross-link with amyloid beta (A\(_\beta\)) and increase risk for both cerebrovascular disease and cognitive impairment, providing an explanation for the elevated risk of dementia in type 2 diabetes individuals\textsuperscript{1,36}. To date, most studies examining associations of glycation with cognitive impairment assess a small number of specific glycation end products\textsuperscript{36,37}. In this glycoproteome-wide study we show that numerous glycation modifications, not measured previously, may contribute to cognitive impairment, further supporting the involvement of glycation in pathological cognitive aging.

Using our new methods, we were able to capture the two types of sugar-adding reactions, glycation and glycosylation. The identification of an unprecedented large quantity of glycation-derived pathological post-translational modifications allowed to locate map numerous changes on the same proteins that were associated with cognitive impairment. The abundance of haptoglobin glycation differed in those with cognitive impairment compared to those without impairment. We found glycation sites on several amino acids on the haptoglobin alpha chain: 131, 141, 136, 153, 251 and 255. These amino acids were located at two different functional sites of the protein: The Serine proteases trypsin domain, located between amino acids 162 and 404, and the SUSHI domain, between amino acids 90 and 147. Such differentiation is not possible when examining proteins, only when examining peptides. These results further emphasize that specific sites at which the glycation modifications are located may be functionally implicated in cognitive impairment.

We also found that many changes in glycosylation were located on the same protein. Moreover, certain glycosylated sites on the same protein differ significantly between those with cognitive impairment and cognitively normal participants, while other glycosylated sites do not differ between the groups. This variation in glycosylation and glycation across multiple sites of a given protein has recently been termed meta heterogeneity\textsuperscript{38} and we confirm here that it is common. The specificity of modified peptide detection will enable us in the future to link modified active sites on a protein with the resulting malfunction of that protein. This demonstrates the potential biological and functional significance to cognition of specific glycoproteoforms, but not others, on the same protein.

This study has several limitations. The IDCD study includes only type 2 diabetes older adults so results from this study cannot be generalized to individuals without type 2 diabetes. Limitations of the study also include the small sample size, which is one of the most important issues in experimental design in clinical proteomic studies\textsuperscript{39}. However, the use of a pilot experiment in proteomic studies was shown to work well when compared with larger studies, providing the samples are representative of the
heterogeneity within a defined population. As in other proteomic-based discovery studies, we did not apply adjustment for multiple comparisons. This study, which to the best of our knowledge is the first global glycoproteome study of its kind on human sera, serves as a basis for research in larger sample sizes and longer follow up for the identification of novel molecular pathways underlying cognitive impairment in type 2 diabetes. Our results corroborate the potential role of glycoproteoforms in type 2 diabetes-related cognitive impairment.

This study has important strengths. It lends support that a novel glycoproteome profiling approach may facilitate large-scale glycoproteome studies to identify a large number of both glycated and glycosylated glycoproteoforms, and the location of site-specific protein modifications. Longitudinal profiling of glycoproteome, in-depth cognitive function and assessment of type 2 diabetes in older adults highlights the potential application of these methods to identify the biology underlying late-life cognitive impairment in type 2 diabetes.

Methods

Participants

Participants were from the ongoing Israel Diabetes and Cognitive Impairment (IDCD) study, a community-based longitudinal cohort study of the chronic conditions of aging in older adults with type 2 diabetes that began in 2010. Participants were from the Maccabi Health Services (MHS), the second largest HMO in Israel, whose exquisitely detailed electronic medical records include diagnoses, blood exams, and medications since 1998.

Ethical considerations: The study was approved by all three IRB committees (Mount Sinai, Sheba Medical Center and Maccabi Health Services). All participants signed informed consent (IRB number 5555-18). All procedures performed were in accordance with the ethical standards of the three IRB committees. Informed consent was obtained from all participants and/or their legal guardians. All participants provided written consent agreeing to annual detailed clinical evaluation. Criteria for enrolment into the IDCD study are: (1) having type 2 diabetes; (2) > 65 years of age, (3) being free of major neurological (e.g., Parkinson's disease, stroke), psychiatric (e.g. schizophrenia) or other diseases (e.g., alcohol or drug abuse) that might affect cognition; (4) having an informant; (5) fluency in Hebrew; (6) living in the area of Tel-Aviv; (7) normal cognition at study entry.

Cognitive assessments

All participants underwent cognitive testing three times 18-months apart, with a battery of 15 tests that were summarized into a composite measure of global cognition by averaging the z-scores from all individual tests. A multidisciplinary diagnostic consensus conference determined the cognitive status at each visit based on all available clinical data as well as the Clinical Dementia Rating scale (CDR). Status included no cognitive impairment, mild cognitive impairment, Alzheimer’s disease dementia, or
dementia from other causes based on published criteria. All participants were initially cognitively normal (CDR=0). The IDCD is an ongoing study and to date has recruited 1211 eligible participants. We randomly selected 8 participants who developed cognitive impairment, defined as a CDR>0.5 (possible dementia or frank dementia) in both follow ups to ensure definite cognitive impairment, and 15 participants who maintained normal cognition over time, i.e. CDR=0 (no dementia) at all three time points. The two groups were matched on age, sex, education, baseline HbA1c and duration of type 2 diabetes.

**Collection of blood samples and clinical characteristics**

Blood samples were obtained in the morning after a 10-h fast. Blood was drawn into serum separator tubes, centrifuged, aliquoted into 0.1 ml vials, and stored at -80 C. Information on demographic (age, gender and years of education) was collected at baseline of the IDCD. Duration of was calculated as the difference between IDCD baseline and the first instance of a diagnosis in Maccabi. HbA1c, systolic and diastolic blood pressure (measured sitting), body mass index (measured as weight in kg divided by the square of height in centimeters), creatinine and cholesterol levels were calculated as the average of all lab results available in Maccabi for each participant until the IDCD study baseline.

**Glycoproteomics analysis**

To identify glycoproteoforms (i.e., the specific glycan with the specific amino acid sequence it binds to) associated with cognitive impairment we performed label free, quantitative glycoproteomics on the sera collected at baseline and at 36 months. The workflow includes proteolytic digestion of the proteins using trypsin, enrichment of glycoproteoforms, LC-MS/MS data acquisition and data processing (Fig. 4).

**Sample preparation**

100ug from each frozen serum sample was digested with trypsin using the S-trap method as detailed in the supplementary material.

**Glycosylation enrichment**

To extract the maximal percentage of glycoproteoforms from the sera samples, glycoproteoforms were enriched by using polyethyleneimine (PEI)-benzoboroxole beads prepared in-house as detailed in the supplementary material.

**Mass spectrometry**

Samples were reconstituted in 15µl 3% Acetonitrile/0.1% formic acid. 1.5µl were loaded on a Symmetry trap column (C18, 180µm*20mm, 5µm, 100A, Waters inc.) followed by a HSS T3 analytical column (C18, 75µm*250mm, 1.8µm, 100A, Waters inc.), mounted on a nanoAcquity running at a flow of 0.35µl/min using a gradient of 4-25% in 125min, followed by 25-40% in 30 minutes. Data was acquired with the Orbitrap Fusion Lumos (Thermo Fisher scientific) running 3sec top-speed data dependent acquisition.
(DDA) method. MS1 scans were performed at 120,000 resolution (at 200 m/z), in 400-1800m/z range. Most abundant ions at charge states 2-8, and at minimum 5e4 intensity were chosen for fragmentation. Precursors were isolated in the quadrupole using 1m/z isolation window. MS2 fragmentation was performed using ETHcD using calibrated charge-dependent parameters with supplemental activation of 15 NCE. Data was acquired at 15,000 resolution (at 200m/z), using a first mass of 120 m/z with standard automatic gain control (AGC) and maximum injection time of 120ms.

**Glycoproteomics data processing and analysis**

Raw data spectra were processed by Byonic software (22) version 4.0.1, using Byonic's Glycoproteoform Search. Data was searched against the UniProt human database (23) with Byonic's common contaminants library appended and a customized reference glycan database containing 84 glycan compositions and 6 most common O-glycans.

Searches were performed using specific cleavage of trypsin with 2 missed cleavages allows, with ETHcD fragmentation. Mass tolerances were set to 10ppm for MS1 and 20ppm for MS2. The following modifications were allowed: fixed carbamidomethylation on C, variable oxidation on M (common1), deamination on NQ (common 1), phosphorylation on STY (rare 1), hex on KW (common 2), protein N-terminal acetylation (rare 1) and peptide N-terminal pyroGlu (rare 1), for a total of two common modifications and one rare modification.

We filtered for glycoproteoforms with Byonic identification quality score of over 150, and quantified them by FlashLFQ (24) using default settings with match-between-runs and normalization. After filtering out non-glycosylated peptides, we summed intensity values of glycoproteoforms with and without Methionine oxidation (+15.9949 Da), asparagine and glutamine deamidation (+0.9840 Da), to avoid doubled quantification of the same peptides. We then filtered out peptides that had an intensity value of less than 30% of at least one of the two study groups.

We were able to identify and quantify glycans with different compositions linked to the same protein site. Fig S3 illustrates the types of modifications that can be found on a single protein. Serum proteins can be modified by two types of sugar-adding reactions: glycation and glycosylation. Our new technology is able to identify both glycation and glycosylation-derived modifications, which can exist within a single protein (Fig. S3A). While only a single glycan can bind to a specific protein site, different glycans can bind to the same protein site, leading to different glycoproteoforms (Fig. S3B).

**Statistical analyses**

The overall goal of these analyses was to examine if there were group differences between the glycoproteomic profiles in individuals with normal cognition throughout the study versus those who developed cognitive impairment. We used Kolmogorov-Smirnov test to examine for group differences in age, sex, HbA1c, duration of type 2 diabetes and number of years of education. Then we examined if there were group differences in the glycoproteoform profiles at baseline. In a further analysis, we
examined if there were group differences in the temporal changes in glycoproteoform abundance by comparing the fold changes between baseline and the 36-months evaluation. We tested for differences in both analyses using Wilcoxon–Mann-Whitney test. Significant p-values were set to 0.05 as done in prior discovery glycoproteomic studies\textsuperscript{40–42,44}.

All statistical analyses were performed in R (version 4.1.1)\textsuperscript{45}.

Biological function enrichment analyses

In the absence of a means for functional glycoproteoform annotation, we used the DAVID protein annotation tool \textsuperscript{46} to group the proteins of the glycoproteoforms based on their biological function.

\textbf{Declarations}

\textbf{Author Contributions Statement}

N.T., D.M., H.L.W and Y.L. researched data, contributed to discussion, and wrote, reviewed, and edited the manuscript. M.S.B and A.S.B reviewed and edited the manuscript. J.U., I.C and R.U. reviewed and edited the manuscript. All authors approved the final version of the manuscript. Y.L. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

\textbf{Additional Information}

This study was funded by the Dr. Marina Nissim foundation, the Leroy Schecter foundation, and NIH grants R01 AG043878, R01 AG051545, R01 AG053446, P50 AG05138 and R01 AG034087.

The authors declare no conflict of interest.

\textbf{Data availability}

The datasets generated during and/or analyzed during the current study are available in the ProteomeExchange, repository, http://www.proteomexchange.org/.

\textbf{References}


**Tables**
Table I  Demographic & clinical data of study cohort. Inter-variance analysis between the parameters of those with cognitive impairment and those without cognitive impairment was performed by Kolmogorov-Smirnov test of homogeneity of variances. There were no significant differences between those with cognitive impairment and those without cognitive impairment in any of the demographic or clinical characteristics. (Mean [SD])

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Developed cognitive impairment</th>
<th>Maintained normal cognition</th>
<th>Kolmogorov-Smirnov test results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>p-value</td>
</tr>
<tr>
<td>Age at baseline (yrs)</td>
<td>75(2.85)</td>
<td>72.14(3.2)</td>
<td>0.51</td>
</tr>
<tr>
<td>Sex (N)</td>
<td>Female 1</td>
<td>4</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Male 7</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Education (yrs)</td>
<td>12.38(3.02)</td>
<td>13.14(3.26)</td>
<td>0.5</td>
</tr>
<tr>
<td>Global Cognition at baseline</td>
<td>Mean=0.33, SD= 0.44</td>
<td>Mean=0.07, SD= 0.56</td>
<td>0.10</td>
</tr>
<tr>
<td>Duration of T2D (yrs)</td>
<td>10.12(3.7)</td>
<td>8.87(4.49)</td>
<td>0.64</td>
</tr>
<tr>
<td>Mean BMI (kg/m2)</td>
<td>26.91(3.25)</td>
<td>27.56(2.86)</td>
<td>0.29</td>
</tr>
<tr>
<td>Mean systolic blood pressure (mmHg)</td>
<td>132.8(8.26)</td>
<td>132.74(6.21)</td>
<td>0.83</td>
</tr>
<tr>
<td>Mean diastolic blood pressure (mmHg)</td>
<td>72.05(3.68)</td>
<td>71.95(4.03)</td>
<td>0.5</td>
</tr>
<tr>
<td>Mean cholesterol (mmol/L)</td>
<td>172.42(21.2)</td>
<td>161.68(22.42)</td>
<td>0.54</td>
</tr>
<tr>
<td>Mean Creatinine (µmol/L)</td>
<td>1.12(0.32)</td>
<td>1.05(0.32)</td>
<td>0.28</td>
</tr>
<tr>
<td>Mean HbA1c (mmol/mol)</td>
<td>6.99(0.55)</td>
<td>6.47(0.5)</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Figures
Figure 1

While most of the 201 proteins had a single glycoproteoform, several proteins had more than 40 glycoproteoforms (bar graph). For example, haptoglobin (HP) had 171 glycoproteoforms. The framed figure shows two of the Haptoglobin glycoproteoforms we found where glycans with different compositions can be attached to the same HP peptide, 'NLFLNHSENNATAK'.
Results of DAVID functional annotation clustering tool. 137 glycoproteoforms from 86 proteins were differentially associated with the group that developed cognitive impairment and the group with stable cognition. Applying the DAVID protein annotation tool to these proteins highlights differences in their functional pathways. X-axis shows the percentage of the proteins that participate in the pathway out of the total proteins. P-value post Benjamini-Hochberg correction of the enrichment is shown on each bar.

A. At baseline there were significant group differences in the abundance of glycoproteoforms from 37 proteins that participated in proteolysis, blood coagulation and immune system related pathways.

B. During three years of follow-up there were group differences in the abundance of glycoproteoforms from 66 proteins between both groups. The main pathways found for these proteins were proteolysis, blood coagulation and immune system related pathways.
Figure 3

Glycosylated glycopeptiforms from five immunoglobulins, IGHG1, IGHG2, IGHG3, IGHG4 and IGHA2, significantly changed between converters and non-converters from baseline to 36 months.

Figure 4
A schematic description of the pipeline used to identify the glycoproteoforms associated with cognitive impairment.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial.docx](#)