Biomarker discovery in diabetes mellitus and lipid metabolism: multi-platform glyco(proteo)mic approaches

Demus, D.A.

English Summary

The core of the research presented in this thesis lies in glycan biomarkers discovery in diabetes, analytical method development, and their clinical implications. Glycan biomarkers have emerged as promising tools for diagnosing, prognosing, and monitoring various diseases, including cancer, autoimmune diseases, and infectious diseases. This thesis focuses on developing and optimizing analytical methods to study glycosylation changes as potential biomarkers in large clinical cohorts of patients with diabetes and related complications.

The research presented aims to address challenges in high-throughput sample preparation of glycosylation analysis, data processing, and statistical analysis. The developed methods are applicable for analysing *N*-glycans, *O*-glycosylated proteins, and absolute fucosylation levels of proteins in blood plasma.

The thesis opens with a general introduction describing types of glycosylation, glyco(proteo)mics analytical approaches, and glycan biomarkers in diabetes (**Chapter 1**). The research part begins with an exploration of *N*-glycan antennary fucosylation as a biomarker for a monogenic type of diabetes, HNF1A-MODY (**Chapter 2**). Individuals with HNF1A-MODY carry variants in the *HNF1A* gene, encoding for a transcription factor HNF1α, which is a master regulator of plasma protein fucosylation. The presence of loss-of-function variants in the *HNF1A* gene leads to upregulation of core fucosylation and downregulation of antennary fucosylation of N-glycans. A new liquid chromatography with tandem mass spectrometry (LC-MS/MS) method was developed to assess fucosylation levels for 320 patient blood plasma samples and evaluate the biomarker's diagnostic performance across different research centres. The results showed a strong correlation of the measured fucosylation levels between the two centres, with correlation coefficients of up to 0.88 for the relevant glycosylation traits. The improved chromatographic separation allowed for the identification of six single glycan traits and a derived antennary fucosylation trait, which effectively differentiated between pathogenic mutation carriers and those with benign or no mutations,

achieving an area under the curve (AUC) of up to 0.94 in the receiver operating characteristic curve analysis.

The subsequent study introduces an enzymatic plate-based assay for assessing α1-3,4 fucosylation levels, aiming to overcome the applicability bottleneck of LC methods in public diagnostic centres (**Chapter 3**). Previous research has demonstrated the effectiveness of the *N*-glycan biomarker in differentiating HNF1A-MODY cases using LC methods. In the current study, a high-throughput exoglycosidase plate-based assay was developed to measure α1-3,4 fucosylation levels in blood plasma samples. This assay was optimised and validated with 1000 clinical samples from a cohort of young-adult onset diabetes patients, including HNF1A-MODY and type 2 diabetes cases. The α1-3,4 fucosylation levels effectively differentiated cases with pathogenic *HNF1A* variants, achieving an AUC value of 0.87. This method was evaluated against previously applied LC methods.

In the second research part, altered apolipoprotein CIII (apo-CIII) *O*-glycosylation profiles linked to increased plasma triglyceride levels in diabetic dyslipidemia are investigated. A highly-automated ultra-high resolution MALDI-FTICR MS method for analysing intact apo-CIII was optimized for large sample cohorts incorporating a chemical oxidation step to reduce methionine oxidation heterogeneity and spectrum complexity (**Chapter 4**). Sinapinic acid matrix minimised sialic acid loss during MALDI measurements, and the MassyTools software was applied for standardised and automated MS data processing and quality control. When applied to 771 plasma samples from individuals without diabetes to access relative levels of apo-CIII glycoforms, the method validated the relationship between apo-CIII glycoforms and lipid biomarkers. The study supports the hypothesis that apo-CIII sialylation influences triglyceride clearance, independent of body mass index.

Subsequently, the proposed MALDI MS-based analytical workflow was applied to a large patient cohort involving 2318 participants from the DiaGene study, including participants with and without type 2 diabetes (**Chapter 5**). This study aimed to unravel

how apo-CIII glycosylation impacts lipid traits and associates with type 2 diabetes prevalence, and to explore the genetic basis of these effects through a genome-wide association study (GWAS). Variants in the *GALNT2* gene, linked to the overexpression of *O*-glycosylating enzyme GALNT2, and the *IFT172* gene were associated with specific apo-CIII glycosylation patterns, high-density lipoprotein (HDL) cholesterol and triglyceride levels. High non-glycosylated apo-CIII (apo-$CIII_{0a}$) levels were associated with increased HDL cholesterol and triglycerides, whereas, disialylated apo-CIII (apo-$CIII_2$) itself was linked to lower triglyceride levels. Replication of these genetic associations with lipid levels was performed in an additional cohort of 5409 individuals from the Diabetes Care System.

The final study investigates associations between apo-CIII glycosylation, genetic variants, and micro- and macrovascular complications of diabetes (retinopathy, nephropathy, neuropathy, cardiovascular disease) in two cohorts: the DiaGene study, n = 1571 and the Hoorn DCS cohort, n = 5409 (**Chapter 6**). Mono-sialylated apo-CIII (apo-$CIII_1$) and disialylated apo-CIII (apo-$CIII_2$) were associated with a reduced and increased risk of retinopathy, respectively. The *GALNT2* gene variant rs4846913, which is associated with lower glycosylated apo-CIII (apo-$CIII_{0a}$) levels, was linked to a decreased prevalence of retinopathy. Higher levels of apo-$CIII_1$ were associated with an increased risk of neuropathy and lower levels of apo-$CIII_{0a}$ were associated with a reduced risk of macrovascular complications.

This thesis concludes with critically evaluating the research findings and their potential clinical implications. Challenges in translating glycan biomarkers into clinical practice are discussed, along with the evaluation of analytical methods and statistical approaches for biopharma industry use.