Sveučilište u Zagrebu

Farmaceutsko-biokemijski fakultet

Dinko Šoić

VISOKOPROTOČNA ANALIZA N-GLIKOZILACIJE KOMPONENTE KOMPLEMENTA C3 KAO BILJEGA ŠEĆERNE BOLESTI TIPA 1

DOKTORSKI RAD

Mentor:

prof. dr. sc. Olga Gornik Kljaić

Zagreb, 2023.

Sveučilište u Zagrebu

Faculty of Pharmacy and Biochemistry

Dinko Šoić

HIGH-THROUGHPUT N-GLYCOSYLATION ANALYSIS OF COMPLEMENT COMPONENT C3 AS BIOMARKER OF TYPE 1 DIABETES

DOCTORAL DISSERTATION

Supervisor:

Prof. Olga Gornik Kljaić, PhD

Zagreb, 2023.

SAŽETAK

Ubrzani razvoj analitičkih tehnika u posljednjih nekoliko desetljeća omogućio je temeljito istraživanje procesa N-glikozilacije u brojnim patofiziološkim stanjima, a uočene promjene u mnogima od njih dovele su do strelovitog rasta interesa za glikobiologiju. Promjene u N-glikozilaciji danas se smatraju vrijednim biljezima raznih bolesti, te se njihova specifičnost pokušava iskoristiti u diferencijalne, dijagnostičke i prognostičke svrhe. Prethodne studije pokazale su da je N-glikanski profil plazmatskih proteina promijenjen u šećernoj bolesti tipa 1 (ŠBT1) i to kako u mladoj populaciji u ranoj fazi bolesti, tako i u odrasloj populaciji s komplikacijama ove bolesti. Jedna od izraženije uočenih promjena bila je ona u zastupljenosti visoko-manoznih glikanskih struktura u plazmatskom N-glikomu. Ove strukture dominantno bi mogle potjecati s komponente komplementa C3, budući da je to glikoprotein s isključivo visoko-manoznim glikanima vezanima na proteinsku okosnicu za kojeg je poznato da doprinosi razvoju ŠBT1 pojačavanjem autoimunih upalnih procesa. Iz tog je razloga u ovom radu razvijena visokoprotočna metoda za analizu N-glikozilacije pojedinih glikozilacijskih mjesta ljudskog C3 proteina upotrebom tekućinske kromatografije spregnute sa spektrometrom masa (LC-MS) temeljena na njegovu obogaćivanju iz plazme pomoću lektinskog afinitetnog medija visokog afiniteta za manozu. Temporalna stabilnost C3 N-glikoma utvrđena je longitudinalnom analizom uzoraka 14 zdravih ispitanika uzorkovanih u tri vremenske točke kroz 10 tjedana. Novorazvijenom metodom zatim je analizirana krvna plazma 61 djece i mladih s novodijagnosticiranih šećernom bolesti tipa 1 i 84 njihove zdrave braće i sestara te su uočene značajne promjene C3 N-glikoma. ŠBT1 povezana je s porastom manje procesuiranih struktura s više manoznih podjedinica na oba N-glikozilacijska mjesta, a regresijski model temeljen na C3 N-glikanima pokazao je snažnu moć u diskriminaciji ispitanika sa ŠBT1 i zdravih kontrola. Nadalje, C3 N-glikozilacija je analizirana u odrasloj populaciji od 189 ispitanika s različitim stupnjem najčešćih komplikacija ŠBT1 – retinopatije i albuminurije. Pokazano je da se C3 N-glikom značajno mijenja u teškoj albuminuriji u ŠBT1, no da je neovisan o samom trajanju bolesti. Retinopatija je pak dovela do promjene samo jedne glikoforme, dok su svi osim jednog C3 glikopeptida značajno povezani s razinama hemoglobina A1c (HbA1c). Predstavljene spoznaje ukazuju na uključenost N-glikozilacije i C3 komponente komplementa u patofiziologiju šećerne bolesti tipa 1 te upućuju na značaj C3 N-glikoma kao potencijalnog dijagnostičkog i prognostičkog biljega ove bolesti i njoj pridruženih komplikacija.

Ključne riječi: N-glikozilacija, N-glikani, glikoproteomika, C3 komponenta komplementa, sustav komplementa, šećerna bolest tipa 1, LC-MS, spektrometrija masa

SUMMARY

Introduction

N-glycosylation is a ubiquitous co- and post-translational modification that enriches the structure and function of proteins, and as such is essential for physiological as well as pathological cellular functions. This process is indirectly encoded by the genome and considerably influenced by environmental factors. Changes in N-glycosylation have been described in different diseases including type 1 diabetes (T1D) and are increasingly considered as biomarkers of ongoing pathological conditions. Previous study demonstrated that the N-glycome of total plasma proteins is changed in children with early-onset T1D compared to their healthy siblings. Furthermore, N-glycosylation was associated with type 1 diabetes complications, as N-glycome of total serum proteins in adult T1D patients with kidney disease was proved to be altered. Among the more pronounced observed associations in both studies were high-mannose N-glycan levels. These structures could predominantly originate from the C3 component of complement, as this protein is exclusively occupied with high-mannose glycans attached to its protein backbone. Besides, it is known that C3 contributes to the development of T1D by enhancing autoimmune inflammatory processes, and the presence of its activated forms in glomeruli and glomerular capillaries of animal models supports its role in diabetic nephropathy. As the most abundant complement protein with plasma concentrations of around 1.2 mg/ml, C3 plays a key role in immune surveillance, so it is no wonder that is also associated with pathophysiological responses. Notwithstanding the fact of involvement of complement and N-glycans in the development of type 1 diabetes, N-glycosylation of C3 complement component is still poorly investigated. One of the reasons for this is the lack of appropriate analytical methods. Although the C3 N-glycan profile is already known, available analytical protocols are only suitable for a limited number of large volume samples. The presence of a high-throughput method for the C3 N-glycome analysis would enable identification of C3 glycoforms associated with the onset of T1D as well as the development of its complications, after which an evaluation of their diagnostic and prognostic potential could be done. All this could lead to discovery of novel risk factors of T1D development and progression, and would result in new knowledge about the mechanism of the disease and the role of glycosylation behind it.

Materials and methods

The first goal of this thesis was to develop a novel cost-effective glycoproteomic 96-well workflow for a high-throughput and site-specific N-glycosylation analysis of human C3 on a glycopeptide level. The glycopeptide level approach guarantees that the changes seen are protein specific and facilitates the value of C3 N-glycoprofiling in type 1 diabetes risk assessment. The analysis is based on the identification and relative quantification of C3 glycopeptides according to their corresponding glycosylation sites. In total, six exclusively high-mannose C3 glycopeptides were found and confirmed in MS/MS analysis that followed, three per each of the two glycosylation sites present. Developed workflow enables the analysis of hundreds of samples in only a couple of days, starting from as low as 10 µl of blood plasma or serum. In order to reduce initial sample complexity, the workflow starts with C3 enrichment using a lectin matrix. Since C3 contains exclusively high-mannose glycans, Concanavalin A (Con A) was chosen as it preferentially recognizes glycans presenting α-mannose-containing cores. Enriched glycoproteins are then digested using endoproteinase Glu-C, and resulting glycopeptides purified from unwanted hydrophobic peptides by solid phase extraction using hydrophilic chromatography interactions (HILIC-SPE). Purified glycopeptides are subsequently analyzed by reversed-phase liquid chromatography coupled to mass spectrometer by electrospray ionization (RP-LC-ESI-MS) technique. Obtained chromatograms are then aligned and glycoproteomic data processed by spectra summing over determined retention time windows. Only signals that meet defined quality control parameters (mass accuracy below 100 ppm, isotopic pattern quality (IPQ) below 25 %, and signal to noise ratio above 8) are eventually summed and used for relative quantification of the glycoform abundance within each glycosylation site. Developed workflow was then applied to several cohorts.

To evaluate the C3 N-glycome stability within a healthy individual under physiological conditions, samples from 14 healthy and age-matched male individuals were measured at three time points: at the beginning of the study, and after 6 and 10 weeks. Temporal stability was assessed by calculating intraindividual and interindividual coefficient of variation (CV) for all C3 glycoforms. Intraindividual CV was measured from longitudinal samples of each subject, while the interindividual CV was computed from all samples within each time point.

In early-onset type 1 diabetes study, plasma samples of 61 newly diagnosed children and adolescents (median age of 10, age range 1-16), as well as 84 unaffected siblings (median age of 11, age range 4-22) derived from the Danish Registry of Childhood and Adolescent Diabetes (DanDiabKids) were analyzed. The Registry contains samples from T1D patients collected

within three months of the disease diagnosis, as well as samples of their healthy siblings. Assessment of T1D association with C3 N-glycome was done using linear mixed modelling, while the evaluation of C3 N-glycome ability to distinguish T1D individuals from their healthy siblings was done using logistic mixed model elastic net regression.

Type 1 diabetes complications study consisted of 189 blood serum samples collected at the Vuk Vrhovac University Clinic for Diabetes, Endocrinology and Metabolic Diseases. Study included type 1 diabetes subjects aged between 18 and 70 years (median age 46 years, age range 18-70), on adequate treatment. Subjects were assessed for common T1D complications (albuminuria and retinopathy) and divided into three categories per complication: without or with initial complication, moderate or severe complication. Assessment of association between C3 N-glycome and various laboratory parameters and T1D complication status was done using general linear modelling, while the ability of C3 N-glycome to differentiate individuals with and without different type 1 diabetes complications was done using logistic elastic net regression.

Results

Developed workflow showed solid inter-day reproducibility with median coefficient of variation being below 12 % for all the glycoforms analysed. C3 N-glycosylation proved to be mostly stable under physiological conditions, with its intraindividual variation being smaller than the variation between individuals, allowing its consideration as a biomarker. Regarding the early-onset T1D study, significant changes of C3 N-glycan profiles were found. Recent-onset type 1 diabetes was associated with an increase in the proportion of unprocessed glycan structures with more mannose units on both C3 N-glycosylation sites: $Glc_1Man_9GlcNAc_2$ on C3.Asn939 glycosylation site and $Man_7GlcNAc_2$ on C3.Asn85 site ($p = 5.43\times10^{-5}$ and $p = 7.82\times10^{-8}$ respectively). A regression model including C3 N-glycans showed notable discriminative power between children with early-onset type 1 diabetes and their healthy siblings with AUC of 0.879.

Regarding the T1D complications study, C3 N-glycome in T1D patients changed significantly in severe stage albuminuria accompanying the disease compared to moderate and initial albuminuria cases, with biggest change being the increase of the monoglucosylated glycoform ($p = 0.002$). C3 N-glycan profile of moderate albuminuria was shown to be fairly similar to that of initial complication. The monoglucoylated form proved to be the only glycoform changed in retinopathy ($p = 0.038$), while smoking status and estimated glomerular filtration rate (eGFR)

in T1D showed no effect on C3 N-glycan profiles. Hypertension status in T1D significantly affected relative abundance of all except one C3 glycoform, with increase in the proportion of glycopeptides with more mannose units evident on both glycosylation sites ($p = 2.23\times10^{-4}$ and $p = 2.71\times10^{-4}$ respectively). The same trend was present with poor glycaemic control: higher HbA1c levels were also associated with glycopeptides with more mannose units on both glycosylation sites. Once again, the monoglucosylated glycoform had the strongest association ($p = 1.97\times10^{-17}$). Lastly, type 1 diabetes duration had no influence on the C3 N-glycome.

Conclusions

A novel high-throughput workflow for detailed N-glycosylation analysis of human C3 was successfully developed and validated. It enables weekly analysis of hundreds of blood plasma or serum samples from as low as 10 µl of starting volume, facilitating the research of C3 N-glycosylation on a large number of samples in future studies of various diseases. As C3 N-glycome proved to be stable in healthy individuals during homeostasis, its potential use as a biomarker is enabled.

Aberrant C3 N-glycan profiles characteristic of the early phase of type 1 diabetes have been identified. Children newly diagnosed with type 1 diabetes were associated with an increase in the proportion of unprocessed glycans, that is structures with more mannose units, on both C3 N-glycosylation sites. A C3 glycan-based discriminative model for recent-onset type 1 diabetes was built and its solid distinguishing power indicates that C3 N-glycome could be valuable in assessment of type 1 diabetes risk in children.

C3 N-glycome was proved to be changed in severe albuminuria accompanying type 1 diabetes, the biggest change being the increase in the proportion of the monoglucosylated glycoform. A significant increase of the same glycoform was also characteristic of non-proliferative retinopathy. C3 N-glycosylation was also shown to be associated with glycaemic control, as measured by HbA1c. Being independent of the disease duration, C3 N-glycome could be a potential novel marker of type 1 diabetes onset, progression and severity.

Keywords

N-glycosylation, N-glycans, glycoproteomics, C3 complement component, complement system, type 1 diabetes, LC-MS, mass spectrometry