Talanta 160 (2016) 614–623

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Classification of congenital disorders of glycosylation based on analysis of transferrin glycopeptides by capillary liquid chromatography-mass spectrometry

Albert Barroso, Estela Giménez, Fernando Benavente*, José Barbosa, Victoria Sanz-Nebot

Department of Chemical Engineering and Analytical Chemistry, University of Barcelona, Diagonal 645, 08028 Barcelona, Spain

ARTICLE INFO

Article history:
Received 30 April 2016
Received in revised form
22 July 2016
Accepted 24 July 2016
Available online 26 July 2016

Keywords:
Congenital disorders of glycosylation
Glycopeptide
Glycoforms
Mass spectrometry
Multivariate data analysis

ABSTRACT

In this work, we describe a multivariate data analysis approach for data exploration and classification of the complex and large data sets generated to study the alteration of human transferrin (Tf) N-glycopeptides in patients with congenital disorders of glycosylation (CDG). Tf from healthy individuals and two types of CDG patients (CDG-I and CDG-II) is purified by immunoextraction from serum samples before trypsin digestion and separation by capillary liquid chromatography mass spectrometry (CapLC-MS). Following a targeted data analysis approach, partial least squares discriminant analysis (PLS-DA) is applied to the relative abundance of Tf glycopeptide glycoforms obtained after integration of the extracted ion chromatograms of the different samples. The performance of PLS-DA for classification of the different samples and for providing a novel insight into Tf glycopeptide glycoforms alteration in CDGs is demonstrated. Only six out of fourteen of the detected glycoforms are enough for an accurate classification. This small glycoform set may be considered a sensitive and specific novel biomarker panel for CDGs.

1. Introduction

Protein glycosylation, which enhances the functional diversity of proteins and provides a highly distinct structure variation, is by far the most common and complex post-translational modification with more than half of all secretory and cellular proteins being glycosylated [1–3]. The lack or attachment of certain monosaccharides to core glycan or branches may result in the alteration of the normal function of the glycoprotein and the resulting transformation of cellular phenotypes is known to be involved in various biological or pathological processes [4].

This is the case of congenital disorders of glycosylation (CDGs), formerly known as carbohydrate-deficient glycoprotein syndromes. CDGs are a family of genetic defects caused by mutations in the genes coding for enzymes involved in the biosynthesis or remodelling of the oligosaccharide moieties of glycoconjugates [5–8]. Depending on the defective biosynthesis step, CDGs can be classified in two subgroups [9,10]. CDG type I (CDG-I) consists in defects in the glycan assembly and in the attachment of glycans to the nascent glycoprotein in the cytosol or the endoplasmic reticulum. This results in the lack of complete N-linked glycans on some glycosylation sites [8]. CDG type II (CDG-II) is the result of abnormal remodelling or processing of the glycan moieties in the Golgi network generally resulting in truncated or structural deficient carbohydrate chains [5,8,11]. Human transferrin (Tf) is a well-known biomarker of CDGs [5,12–14]. Tf (~80 kDa, ~5.8% carbohydrates) is an iron-binding serum transport glycoprotein with one O-glycan (one hexose unit at Ser32), and two complex N-glycans at Asn413 and Asn611 [15]. Due to the microheterogeneity associated with the complex N-glycans, Tf exists as a mixture of glycoforms which differ in composition, structure and charge. The main glycoform of Tf is the tetrasialoform (S4, 85% of the total amount of Tf in healthy individuals) that shows two disialylated biantennary glycans (2Ant2SiA+2Ant2SiA; antennae (Ant) and sialic acid (SiA)=N-acetylneuraminic acid). Other sialoforms such as pentasialo (S5, 2Ant/2SiA+3Ant/3SiA), hexasialo (S6, 3Ant/3SiA+3Ant/3SiA), trisialo (S3, 2Ant/2SiA+2Ant/1SiA) and disialo (S2, 2Ant/2SiA) are present at much lower concentrations in healthy individuals (S1, 2Ant/1SiA, and S0, no glycosylation, are almost inexistent) [16–18]. With regard to Tf glycosylation in CDG patients, there is a difference between CDG-I and CDG-II. CDG-I is mostly characterised by increased S2 and S0 and decreased S4 while CDG-II mostly shows increased S3, S1 and/or S0 and decreased S4 [8]. Nowadays, the diagnosis of CDGs is mainly based upon the glycoform pattern observed for intact Tf by

Abbreviations: CapLC-MS, capillary liquid chromatography mass spectrometry; Tf, human transferrin; CDG, congenital disorders of glycosylation; EIC, extracted ion chromatogram; PLS-DA, partial least squares discriminant analysis

* Corresponding author.
E-mail address: fbenavente@ub.edu (F. Benavente).

http://dx.doi.org/10.1016/j.talanta.2016.07.055
0039-9140/

isoelectric focusing (IEF) [7,19,20]. However, this method does not allow the unequivocal identification of Tf glycoforms, often lead to false negative results (25% of the identified CDG may show a normal profile of intact Tf by IEF) and sometimes it is difficult to discriminate between CDG-I and CDG-II, or from other glycosylation defects (galactosemia, fructose intolerance, alcohol abuse, etc.) [1,7]. In addition, the analysis of intact Tf is not enough to deeply investigate the mechanism altering glycosylation.

High performance separation techniques coupled to mass spectrometry (MS) are the most reliable way to obtain structural information about protein glycosylation as they allow fast and high sensitivity profiling and accurate characterisation of heterogeneous glycan structures [21–25]. In this regard, the enzymatic digestion of the glycoprotein and the detection of certain glycomarkers of lower molecular mass in the digests (glycopeptides or glycans) is a powerful alternative to detect glycoprotein glycoforms by MS instead of intact glycoproteins, whose sensitivity in MS tends to be lower [26–30]. Furthermore, the analysis of glycopeptides obtained after protein digestion provides information about the structure and composition of the glycans, as well as about the glycosylation sites and their degree of occupancy [31,32]. However, the analysis of glycopeptides of an enzymatic protein digest is a challenge, due to the complexity of the digests, which are mixtures of peptides and glycopeptides, and the microheterogeneity of the glycopeptides, which present several glycoforms [33,34]. As a result, very often, experiments involving the analysis of glycoconjugates result in complex and massive data sets which are easier to interpret using multivariate data analysis methods [35].

In this preliminary study, Tf tryptic digests from healthy individuals and CDG-I and CDG-II patients are analysed to investigate glycosylation alteration. Tf is purified by immunoextraction from serum samples before trypsin digestion and separation by capillary liquid chromatography mass spectrometry (CapLC-MS). Following a targeted data analysis approach, partial least squares discriminant analysis (PLS-DA) is applied to the relative abundance of Tf glycopeptide glycoforms obtained after integration of the extracted ion chromatograms of the different samples. The performance of PLS-DA for classification of the different samples and for providing a novel insight into how differently both N-glycosylation points of Tf can be affected in CDGs is demonstrated.

2. Materials and methods

2.1. Chemicals

All chemicals used in the preparation of buffers and solutions were of analytical reagent grade. Isopropanol (iPrOH), hydrochloric acid (HCl, glacial), formic acid (HFor, 98–100%) and glycine ($\geq$ 99.7%) were supplied by Merck (Darmstadt, Germany). CNBr-activated-Sepharose 4B was provided by GE Healthcare (Waukesha, WI, USA). Sodium chloride (NaCl, $\geq$ 99.5%), DL-Dithiothreitol (DTT, $\geq$ 99%), iodoacetamide (IAA), ammonium hydrogencarbonate, water (LC-MS grade) and acetonitrile (LC–MS grade) were supplied by Sigma-Aldrich (Madrid, Spain) and Tris(hydroxymethyl) aminomethane (TRIS, $\geq$ 99.5%) by J.T. Baker (Deventer, Holland). Trypsin (Sequencing grade modified) was provided by Promega (Madison, WI, USA). RapiGest® from Waters (Bedford, MA, USA) was used to facilitate the enzymatic digestion. Goat polyclonal antibody against Tf (immunogen affinity purified) was purchased from Abcam (Cambridge, UK). ESI low concentration (ESI-L) tuning mix was supplied by Agilent Technologies (Waldbronn, Germany) for tuning and calibration of the TOF mass spectrometer.

2.2. Tf standard and serum samples

Tf standard was purchased from Sigma-Aldrich (Madrid, Spain). Tf standards were prepared at 1500 μg mL^{-1} ($\sim$19 μM) in 50 mM NH_4HCO_3 buffer (pH 7.9).

A commercial lyophilised pool of sera with a non-pathological profile (control, SERODOS® plus) was purchased from Bio-Rad (Hercules, CA, USA). Serum control samples from young-adult healthy volunteers were obtained in the Department of Analytical Chemistry of the University of Barcelona [36]. Briefly, venous blood was collected in 9 mL Vacuette tubes (Greiner Bio-One, Frickenhausen, Germany) with Z serum separation clot activator, and then allowed to coagulate by leaving it undisturbed at room temperature for 9 h. Afterwards, the clot was kept at 4 °C for 12–16 h to improve the clot retraction. The supernatant serum was subsequently separated from the clot with a Pasteur pipette and centrifuged at 1200xg for 20 min at 4 °C. Clear serum was then separated and aliquoted to store in a freezer at −20 °C when not in use. Positive CDG-I and CDG-II serum samples from infants and young-adults were kindly provided by the Balagué Center S.A. (Barcelona, Spain), the Institute of Clinical Biochemistry (Hospital Clínic, University of Barcelona) and the Center for the Diagnosis of Molecular Diseases (Autonomous University of Madrid). The assay was approved by the Ethics Committee of the University of Barcelona.

In the last years, a novel nomenclature for the CDGs has been proposed, which connects the name of the defective gene followed by a common -CDG suffix [37]. Nine healthy controls (1–9), five CDG-I (PGM1-CDG (10); DPM1-CDG (11); RFT1-CDG (12); DPAGT1-CDG (13); PMM2-CDG (14)) and five CDG-II (ATP6V0A2-CDG (15); $CDGII_x$ (16); $CDGII_x$(sepsis) (17) and two samples of COG8-CDG (18–19)) samples were analysed for the calibration. The validation set contained nine samples (i–ix): three healthy controls (i–iii), three CDG-I (PMM2-CDG (iv), DPAGT1-CDG (v) and DPM1-CDG (vi)) and three CDG-II (ATP6V0A2-CDG (vii), CDGIIx (viii) and CDGIIx (sepsis) (ix)). The assay was approved by the Ethics Committee of the University of Barcelona.

2.3. Purification and digestion of Tf

Tf was purified from serum by immunoaffinity chromatography with a cyanogen-bromide sepharose column as described in a previous work [33]. Briefly, 50 μL of serum were diluted 1:4 in 10 mM Tris-HCl (pH 7.6). Before passing the sample through the column, a conditioning step was carried out consisting in two washes of 1 mL of 10 mM Tris-HCl (pH 7.6). Afterwards, the diluted serum was passed through the column ten times. After washing with 400 μL of 10 mM Tris-HCl and 0.5 M NaCl (pH 7.6), retained Tf was eluted with 250 μL of 100 mM glycine-HCl (pH 2.5). Eluted Tf was immediately neutralised with 0.5 M Tris pH $\sim$11. Afterwards, the buffer was exchanged for 50 mM NH_4HCO_3 buffer by ultracentrifugation using Microcon YM-10 filters (M_r cut-off 10,000, Millipore, Bedford, MA, USA).

Tf standards and Tf purified from serum samples were reduced, alkylated and immediately subjected to enzymatic digestion in the presence of RapiGest® as explained in [33]. Briefly, 1.90 μL of 0.5 M DTT in 50 mM NH_4HCO_3 buffer was added to an aliquot of 50 μL of standard or purified Tf with 0.1% (w/v) RapiGest®. The mixture was incubated in a TS-100 Thermo-Shaker digester (Bio-San, Riga, Latvia) at 56 °C for 30 min and then alkylated with 50 mM IAA for 30 min at room temperature in the dark (5.25 μL of 0.73 M IAA). Excess reagent was removed by ultracentrifugation with Microcon YM-10 filters, washing 3 times with NH_4HCO_3 buffer. The final residue was reconstituted to the initial volume (50 μL) with NH_4HCO_3 buffer with 0.1% (w/v) RapiGest®. Trypsin in an enzyme to protein ratio of 1:40 (m/m) (considering that Tf concentration is

Please cite this article as: A. Barroso, et al., Talanta (2016), http://dx.doi.org/10.1016/j.talanta.2016.07.055