Manuscript Details

Manuscript number	CHROMB_2019_1168_R1
Title	Comparative analysis of the human serum N-glycome in lung cancer, COPD and their comorbidity using capillary electrophoresis
Article type	Full Length Article

Abstract

Lung cancer (LC) and chronic obstructive pulmonary disease (COPD) are prevalent ailments with a great challenge to distinguish them based on symptoms only. Since they require different treatments, it is important to find non-invasive methods capable to readily diagnose them. Moreover, COPD increases the risk of lung cancer development, leading to their comorbidity. In this study the N-glycosylation profile of pooled human serum samples (90 patients each) from lung cancer, COPD and comorbidity (LC+COPD) patients were investigated in comparison to healthy individuals (control) by capillary gel electrophoresis with high sensitivity laser-induced fluorescence detection. Sixty-one N-glycan structures were identified in the pooled control human serum sample and the profile was compared to pooled lung cancer, COPD and comorbidity of COPD with lung cancer patient samples. It is important to note that no other sugar structures were detected in any of the patient groups, only quantitative differences were observed. Based on this comparative exercise, a panel of 13 N-glycan structures were identified as potential glycobiomarkers to reveal significant changes (>33% in relative peak areas) between the pathological and control samples. In addition to N-glycan profile changes, alterations in the individual N-glycan subclasses, such as total fucosylation, degree of sialylation and branching may also hold important glycobiomarker value.

Keywords	human serum; N-linked glycan; capillary electrophoresis; pulmonary diseases
Taxonomy	Biomarker Research, Capillary Electrophoresis, Glycomics Analysis

Corresponding Author	Brigitta Meszaros
Corresponding Author's Institution	Horvath Csaba Memorial Laboratory of Bioseparation Sciences Research Center for Molecular Medicine Doctoral School of Molecular Medicine Faculty of Medicine University of Debrecen Debrecen Hungary
Order of Authors	Brigitta Meszaros, Gabor Jarvas, Anna Farkas, Marton Szigeti, Zsuzsanna Kovacs, Renata Kun, Miklos Szabo, Eszter Csanky, Andras Guttman
Suggested reviewers	Jana Krenkova, Julia Khandurina, Marcell Olajos, Peter Mizsey

Submission Files Included in this PDF

File Name [File Type]

JCB-2 cover letter.docx [Cover Letter]

Reply to Reviews.docx [Response to Reviewers]

CHROMB_2019_1168R1_with changes.docx [Revised Manuscript with Changes Marked]

Highlights_meszaros_etal.pdf [Highlights]

Abstract_meszaros_etal.pdf [Abstract]

CHROMB_2019_1168R1_revised.docx [Manuscript File]

coi_signed_meszaros_etal.pdf [Conflict of Interest]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Professor Govert Somsen
Editor
Journal of Chromatography B

Debrecen, Oct 14, 2019

Dear Govert :

Please find enclosed our revised manuscript #CHROMB_2019_1168R1, titled as "Comparative analysis of the human serum N-glycome in lung cancer, COPD and their comorbidity using capillary electrophoresis" by Brigitta Meszaros et al that has been corrected by following all recommendations of the Editor and the reviewers along with our detailed replies to the issues raised.

We would like to emphasize that our work presented in this paper was a pilot study focused on human serum N-glycosylation analysis of malignant and chronic pulmonary diseases by capillary electrophoresis. In this pilot study, we used pooled serum samples (90 each) of healthy control, lung cancer, COPD and their comorbidity patients to work out the sample preparation and analysis details. 61 N-glycan structures were identified in the pooled control human serum and compared to pooled LC, COPD and COPD with LC patient samples. Based on our results, we have identified a panel of 13 N-glycan structures as potential glycobiomarkers to reveal significant changes between the pathological and control samples.

In this paper we also introduced some modifications of existing vendor methodologies, which were optimized for purified IgG analysis and not for human serum samples. Along these changes were the temperature adjusted denaturation step to prevent precipitation of the more complex human serum samples and the longer endoglycosidase incubation time to assure complete deglycosylation of the complex biological samples.

The suggested approach readily supports development of early diagnostic tools for lung cancer, COPD and their comorbidity. Therefore, our results should be appealing not only for specialists in capillary electrophoresis, but a wider separation science community. Currently we are in the process of collecting 300 samples from each the disease groups and the resulting 1200 samples will be individually analyzed with the method developed during this pilot study.

Your kind attention to this matter is greatly appreciated.

Sincerely yours,

Prof. András Guttman
Member of the Hungarian Academy of Sciences

Prof Govert W. Somsen
Editor
Journal of Chromatography B

Debrecen, Oct 14, 2019

Dear Govert:
Please find enclosed our revised manuscript # CHROMB_2019_1168R1, titled as: "Comparative analysis of the human serum N-glycome in lung cancer, COPD and their comorbidity using capillary electrophoresis" by Brigitta Meszaros et al that has been corrected by following all recommendations of the Editor and the reviewers along with our detailed replies to the issues raised.

Reviewer 1: We thank the reviewer for the instructive comments that we complied with in our revised version

The presented manuscript reports a comparative study of the human serum N-glycome in various lung diseases using CE. The work is quite clearly presented with adequate references cited. The main comment is related to the informative value of the analysis of pooled samples because some data are inconsistent. A better explanation of the unclear results should be provided. For example, very high increments of tetrasialylated and mono-antennary in the case of LC and COPD (Figure 3), however, the smallest increments in their comorbidity (similarly shown in Table 2). A better approach should be the analysis of individual samples with following statistical analysis, as suggested by the authors in the Conclusion.
REPLY: Considering the reviewer's comment, we added an LC vs COPD column to Table 2 to clarify the issue. Sorry for not adequate in regard to Figure 3, per the reviewer' suggestion, we also put more emphasis in the text that it shows changes in respect to the control only. The comorbidity data was derived from patient samples suffering both diseases simultaneously, so their N-glycome have changed differently that of the addition of LC + COPD would have resulted. To avoid misinterpretation, we changed the LC+COPD terminology for LC with COPD. Also, we would like to emphasize, that this report is about the preliminary data analysis (pilot study) working out the sample preparation and analysis details. We are currently progressing with the measurements of 300 samples each of LC, COPD and LC with COPD and 300 control, thus 1200 samples that is a major undertaking and will take at least another 6-9 months to accomplish.

Minor comment: the style of the references should be corrected.
REPLY: We appreciate the comment and the references have been corrected accordingly.

-**Reviewer 2.** We are grateful for the great comments of the reviewer and addressed the issues raised in the revised version

The authors have provided a study that examines the *N*-glycan profile of a pool of healthy volunteers, pooled patient groups with COPD, lung cancer, COPD and lung cancer. While the study is pleasantly written, the reviewer feels that the performed study may not be suitable for the Journal of Chromatography B. This, as the authors apply a published and commercial available analytical method (Fast Glycan kit with CE-LIF) with little to no modifications. Therefore, the major element of novelty in this case is focused on the application of studying glycomic differences between the different pooled groups. While the subject is relevant in the field of glyco(proteo)mics, the manuscript would maybe be better suited for a more specialized journal.

<u>REPLY</u>: We greatly appreciate the reviewer's comment and realized that we did not emphasize our important methodological developments in the experimental section. Since the Fast Glycan Sample Preparation and Analysis protocol was optimized for purified IgG samples, we had to make important changes to avoid possible precipitation issues with the more complex serum samples, thus, a modified temperature adjusted denaturation protocol was used as added to the Experimental section. Also, the PNGase F digestion time was increased from 20 to 60 min to accommodate the more complex glycoprotein mixture in the human serum samples. Finally, we strongly feel that this paper is suitable for J.Chrom B - Biomedical Applications, as it is focused on a new biomedial application with the associated optimized methodologies presented.

The reviewer has some suggestions before resubmitting this paper:

- The research studies the N-linked glycosylation in serum from different patient pools. The authors mention the fact that the decision was made for studying sample pools rather than individual samples as they were afraid for information loss of analytes that might be under the detection threshold, however, it seems a rather invalid argument as especially by pooling material the high abundant species could get more abundant in the pooled sample as minor abundant species might not be present in all the patient samples. For this reason, it remains unclear for the reviewer why it was chosen to go for pools rather than for individual analysis of the patient samples. Additionally, to improve the impact of the study, the reviewer would suggest to perform additional experiments studying the individual samples rather than comparing the pooled glycan profiles, as the authors already have 90 samples per group available of which only 2 μL is needed for the analysis (e.g. combine this study with the suggested study in the conclusions section or perform a validation study when the 300 samples per group are collected and potential biomarkers are identified in the study with the smaller sample size). This is especially encouraged, as the described assay (Fast Glycan kit, the analytical platform (CE-LIF) and the GUcal software) is an excellent platform to study larger sample sizes. Moreover, with the information obtained from the individual samples an overview can be obtained of the biological variation within each group and to examine if the major differences found between the different pools can also be observed when the biological variation of each group is taken into account.

<u>REPLY</u>: We greatly appreciate the reviewer's point of view on this part and emphasize that it was a pilot study to validate the modified methods used for the patient serum samples. We are currently proceeding with the detailed study of 300 samples each (LC, COPD, LC with COPD and control, 1200 samples total) that will take us for another 6-9 months to finish with the analysis and validation using the modified sample preparation and analysis method described in this paper.

- When technical replicates are performed, next to relative abundance, information should be provided such as standard deviation and relative standard deviation as this provides insight in the technical variation of the assay (for instance this information should be added to Table 1 and Supplementary Table 1).

<u>REPLY</u>: In strong agreement with the reviewer, all requested information with SD and RSD values have been added to the revised manuscript.

- Please use the IUPAC regulations for writing letter symbols in italic when they are locants in chemical-compound names indicating attachments to heteroatoms

(https://www.iupac.org/cms/wp-content/uploads/2016/01/ICTNS-On-the-use-of-italic-and-roman-fonts-for-symbols-in-scientific-text.pdf) (e.g. *N*-glycans).

REPLY: Thanks for noticing this issue that has been corrected in the revised version.

Comparative analysis of the human serum ~~N-~~ *N*-glycome in lung cancer, COPD and their comorbidity using capillary electrophoresis

Brigitta Mészáros[1], Gábor Járvás[1,2], Anna Farkas[1], Márton Szigeti[1,2], Zsuzsanna Kovács[1], Renáta Kun[1,3], Miklós Szabó[3], Eszter Csánky[3] and András Guttman[1,2]

[1]Horváth Csaba Memorial Laboratory of Bioseparation Sciences, Research Center for Molecular Medicine, Doctoral School of Molecular Medicine, Faculty of Medicine, University of Debrecen, Debrecen, Hungary

[2]Research Institute of Biomolecular and Chemical Engineering, University of Pannonia, Veszprem, Hungary

[3]Department of Pulmonology, Semmelweis Hospital, Miskolc, Hungary

Keywords: human serum, ~~N-~~ *N*-linked glycan, capillary electrophoresis, pulmonary diseases

Abbreviations:

COPD – chronic obstructive pulmonary disease
LC – lung cancer
DP – degree of polymerization
APTS – 8-aminopyrene-1,3,6-trisulfonic acid
GU – glucose unit
IgG – immunoglobulin G
IgA – immunoglobulin A
AAT – α~~alpha~~-1-antitrypsin
HP – haptoglobin
TR – transferrin
DTT – dithiothreitol
RNase B – ribonuclease B
RFU –relative fluorescence unit

Abstract

Lung cancer (LC) and chronic obstructive pulmonary disease (COPD) are prevalent ailments with a great challenge to distinguish them based on symptoms only. Since they require different treatments, it is important to find non-invasive methods capable to readily diagnose them. Moreover, COPD increases the risk of lung cancer development, leading to their comorbidity. In this pilot study the ~~N-~~ *N*-glycosylation profile of pooled human serum samples (90 patients each)

from lung cancer, COPD and comorbidity (LC with +COPD) patients were investigated in comparison to healthy individuals (control) by capillary gel electrophoresis with high sensitivity laser-induced fluorescence detection. Sample preparation was optimized for human serum samples introducing a new temperature adjusted denaturation protocol to prevent precipitation and increased endoglycosidase digestion time to assure complete removal of the *N*-linked carbohydrates. Sixty-one N- *N*-glycan structures were identified in the pooled control human serum sample and the profile was compared to pooled lung cancer, COPD and comorbidity of COPD with lung cancer patient samples. One important findings was It is important to note that no other sugar structures were detected in any of the patient groups, only quantitative differences were observed. Based on this comparative exercise, a panel of 13 N- *N*-glycan structures were identified as potential glycobiomarkers to reveal significant changes (>33% in relative peak areas) between the pathological and control samples. In addition to N- *N*-glycan profile changes, alterations in the individual N- *N*-glycan subclasses, such as total fucosylation, degree of sialylation and branching may also hold important glycobiomarker values.

Introduction

Lung cancer (LC) is one of the predominantly diagnosed cancers (11.6% of the total cases) and was the main cause of cancer deaths (18.4% of the total cancer deaths) in 2018 in both genders combined [1]. Another prevalent lung disease is Chronic Obstructive Pulmonary Disease (COPD), mostly caused by cigarette smoking, representing a great but certainly preventable risk factor [2]. COPD has recently become in the spotlight due to its elevating incidence rate, morbidity, mortality and increasing the risk of developing lung cancer, in many cases creating formidable challenges for the global healthcare systems [3]. Simultaneous diagnoses of LC and COPD (i.e., comorbidity) represent poor prognosis [2], thus treatments should be planned accordingly [2] [4]. The high

mortality rate of COPD and lung cancer could be due to the asymptomatic property at their early stages, and the lack of appropriate distinguishing diagnostic tools. While biopsy can differentiate malignant (LC) and benign (inflammation) lesions that imaging techniques are not capable in a reliable diagnostic manner, it is an invasive process with a number of risk factors including infections, pneumonia, pneumothorax, bleeding and haemoptysis, just to list the most frequent ones [5]. Furthermore, another potential drawback of biopsy is the risk of tumor spreading, i.e., during the biopsy process, possible metastasis initiating circulating tumor cells may get away from the primary tumor [6, 7]. Therefore, there is an urgent need to develop non-invasive molecular diagnostic tools capable of predicting the presence and prognosis of the actual disease (LC, COPD or ~~both~~comorbidity) with adequate specificity and sensitivity.

Ever-growing evidence shows the importance of the ~~N-~~ *N*-glycosylation of proteins in biological systems, demonstrating that this post-translational modification is as essential as the polypeptide backbone itself, ~~since they~~ playing significant roles in forming their higher order structure, biochemical properties and function [8]. Suzuki et al. [9] studied the importance and roles of ~~N-~~ *N*-glycosylation in COPD and found that the reduction of FUT8 activity has close relation with the progression of the disease. Phillips and coworkers [10] reviewed the glycosylation aspects of lung cancer and the mechanisms of post-synthetic glycan modification during malignant transformation suggesting promising biomarkers and therapeutic possibilities based on their ~~N-~~ *N*-glycosylation alterations. The significance of these changes revealed that identifying ~~N-~~ *N*-glycan biomarkers as potential early detection markers and monitoring the treatment of lung diseases [2] can be of high importance.

The human serum contains a plethora of proteins with the majority of them glycosylated [11]. Protein glycosylation is the consequence of very complex biochemical processes, regulated by a number of glycosidases and glycosyltransferases [9] resulting in diverse~~,~~ but ~~characterized~~ protein specific glycan profiles, which affect several cellular properties such as signaling, adhesion, motility and half-life, just to mention ~~some~~ a few important ones [12]. Genetic and environmental factors can both affect the activity and level of this glycosylation machinery that may lead to altered glycan structures, possibly specific for pathological changes. Comprehensive analysis of complex carbohydrates requires high sensitivity methods with enhanced resolution because of their ~~low concentration and~~ great structural diversity. The most frequently applied techniques for ~~N-~~ *N*-glycan analysis are high-performance liquid chromatography (HPLC), capillary electrophoresis (CE), mass spectrometry (MS) and NMR, or the combination of those such as LC-MS or CE-MS [13].

In several cases, ~~N-~~ *N*-glycosylation alterations affect the high abundant acute phase proteins, including immunoglobulin G (IgG) [14, 15], immunoglobulin A (IgA) [11, 16, 17], alpha-1-antitrypsin (AAT) [18], haptoglobin (HP) [19, 20] and transferrin (TR) [21, 22]. Pavić et al [15] analyzed the human plasma and its IgG subset of COPD patients and healthy controls by UPLC and provided new insights into plasma protein and IgG ~~N-~~ *N*-glycome changes. The plasma protein ~~N-~~ *N*-glycome significantly decreased in low branched type and increased in more complex glycan structures in COPD patients. Ito et al. investigated the ~~N-~~ *N*-glycans of a lung-specific protein, surfactant protein D, using matrix-assisted laser desorption/ionization quadrupole ion trap time-of-flight mass spectrometry and found that the fucosylation level was greatly elevated in COPD patients compared to controls [23]. Rudd and coworkers analyzed the serum of lung cancer patients

and healthy controls by high performance hydrophilic interaction liquid chromatography (HILIC) and anion exchange chromatography [24]., and They found that the level of bi- (A2), tri- (A3) and tetra-antennary (A4) glycans were significantly increased in lung cancer compared to healthy controls and a reduction was observed in the amounts of core-fucosylated bi-antennary structures. In addition to total human serum, one of the abundant proteins, haptoglobin was analyzed and similar alterations were found in both sample types. Váradi et al. also studied the human serum haptoglobin in lung diseases and emphasized the importance of measuring the core- to arm-fucosylation ratios [20]. The slight decrease in the total fucosylation level of haptoglobin in the serum of COPD and pneumonia patients compared to the control group was the result of a significant decrease in arm-fucosylation and a slight increase in core-fucosylation. Elevated amounts of core- (FA4G4) and antennary-fucosylated tetra-antennary glycans (A4FG4) in haptoglobin were also observed comparing lung cancer and COPD patient groups. Ruhaak et al. investigated the N- *N*-glycans of several high abundance glycoproteins enriched by affinity capture from plasma samples of lung adenocarcinoma patients as well as healthy controls and analyzed with nano-HPLC-chip-TOF-MS to search for lung cancer biomarkers [25]. They found that, while the N- *N*-glycan profiles of blood-derived glycoproteins could be adequate biomarkers for lung cancer, protein enrichment did not improve specificity, but made the method more complicated. Liang et al. [18] examined different types of pulmonary diseases including lung adenocarcinoma, squamous cell lung cancer, small-cell lung cancer as well as several benign ones, such as pneumonia, pulmonary nodule and tuberculous pleuritis and suggested that the N- *N*-glycosylation patterns of α-1-antitrypsin could be a potential lung cancer biomarker.

In this paper we report on ~~the investigationa~~ pilot study of the ~~N-~~ *N*-glycosylation profiles of pooled human serum samples from patients with COPD, lung cancer and their comorbidities (COPD with ~~+~~LC) compared to control healthy subjects (pool of 90 patients each) using capillary electrophoresis with laser-induced fluorescent detection (CE-LIF). A new temperature adjusted denaturation protocol was used to prevent precipitation and the endoglycosidase digestion time was increased to assure full removal of the *N*-linked oligosaccharides. Sixty-one ~~N-~~ *N*-glycans were identified in the pooled, control human serum samples based on data in publicly available GU databases and exoglycosidase digestion based carbohydrate sequencing. The relative peak areas of the identified ~~N-~~ *N*-glycans were used in a comparative quantitative evaluation to find a potential preliminary glycobiomarker panel to differentiate lung cancer, COPD and their comorbidity from each other and from the control.

Materials and methods

Chemicals and reagents

Water (HPLC grade), acetonitrile, sodium cyanoborohydride (1 M in THF), glycerol, ribonuclease B, immunoglobulin G, α~~alpha~~-1-antitrypsin and DTT (dithiothreitol) were obtained from Sigma Aldrich (St. Louis, MO, USA). SDS (sodium dodecyl sulfate) and Nonidet P-40 were from VWR (Radnor, PA, USA). The Fast Glycan Labeling and Analysis Kit was from SCIEX (Brea, CA, USA) including the bracketing standards of maltose (DP2) and maltopentadecaose (DP15). The exoglycosidase enzymes of Sialidase A (*Arthrobacter ureafaciens*), β-Galactosidase (*Jack bean*) and β-N-Acetyl Hexosaminidase (*Jack bean*) were from ProZyme (Hayward, CA, USA). The endoglycosidase PNGase F was from Asparia Glycomics, (San Sebastian, Spain).

Sample preparation

All serum samples were collected with the appropriate Ethical Permissions (approval number: 23580-1/2015/EKU (0180/15)) and Informed Patient Consents at the Department of Pulmonology in the Semmelweis Hospital (Miskolc, Hungary). For this pilot study, ~~T~~the samples were pooled in order to minimize information loss of species below the detection threshold and improve efficiency following the method described in [26]. Serum samples from 90 healthy individuals (control), 90 lung cancer patients, 90 COPD patients and 90 patients with comorbidity of COPD with lung cancer were separately pooled.

Preparation of human serum samples included denaturation, glycan release, fluorophore labeling and magnetic bead mediated cleanup. First, 2 µL of serum samples were diluted with HPLC grade water to 10 µL. Since the Fast Glycan Sample Preparation and Analysis protocol was optimized for purified IgG samples, to avoid possible precipitation issues with the more complex serum samples, a modified denaturation protocol was used by ~~, followed by denaturation~~ adding 5 µL denaturation solution ~~at~~ and applying 40°C for 10 min followed by 70°C for 10 min ~~by adding 5 µL denaturation solution of the Fast Glycan kit (SCIEX)~~. The ~~G~~glycan release process was also modified and performed by the addition of 1.0 µL of PNGase F enzyme (200 mU) to the reaction mixture and incubated at 60°C for 1 hour instead of 20 min, to ensure complete ~~deglycosylation~~removal of the serum *N*-glycome. The endoglycosidase digestion reaction was stopped by the addition of 120 µL of ice-cold acetonitrile. This was followed by drying the reaction mixture under reduced pressure at 60°C for 1 hour in a SpeedVac. The dried samples were reconstituted in the labeling solution containing 4.0 µL of 40 mM 8-aminopyrene-1,3,6-trisulfonic acid (APTS) in 20% acetic acid, 2.0 µL of $NaBH_3CN$ (1 M in THF) and 4 µL 20% acetic acid. The reaction mixture was incubated in a heating block using ~~our earlier published~~a modified

evaporative labeling protocol with closed vial cap at 50°C for 60 min, followed with open cap at 55°C for 80 minutes [27]. After the labeling step, the samples were purified by magnetic beads following the Fast Glycan Sample Preparation and Analysis protocol and analyzed by CE-LIF. Exoglycosidase digestions wereas performed by consecutive additions of sialidase A to remove all α 2-3,6,8-linked sialic acids, Jack bean galactosidase to remove β1-4,6-linked galactose residues and Jack bean hexosaminidase to remove the β1-2,4,6-linked ~~N-~~ *N*-acetyl-glucosamines by respective overnight incubations at 37°C as described earlier in [28].

Capillary electrophoresis

Capillary electrophoresis analyses with laser induced fluorescent detection (CE-LIF) were performed using a PA800 Plus Pharmaceutical Analysis System (SCIEX). All CE measurements were accomplished in 40 cm effective length (50 cm total length), 50 µm ID bare fused silica capillaries filled with the HR-NCHO separation gel buffer (SCIEX). 30 kV electric potential was applied during the separation step in reversed polarity mode (cathode at the injection side, anode at the detection side) at 30°C. To increase detection sensitivity and reproducibility, a~~A~~ three-stage sample injection procedure was used: Step 1) 1.0 psi for 5.0 sec water pre-injection, Step 2) 3.0 kV for 3.0 sec sample injection and Step 3) 2.0 kV for 2.0 sec bracketing standard injection. This latter was used for high precision GU value determination by the GUcal software (www.GUcal.hu) [29]. Data collection and analysis were implemented by the 32Karat (version 10.1) software package (SCIEX). Relative percentage area values of the separated peaks were calculated by the PeakFit v4.12 Software (SeaSolve Software Inc., San Jose, CA).

Results and Discussion

In this pilot study, a new sample preparation method was applied to accommodate the high complexity of the human serum samples including temperature adjusted denaturation and extended glycan release and evaporative labeling protocol. eCapillary electrophoresis with - laser induced fluorescence detection was used was applied to analyze and compare the N- *N*-glycosylation patterns of pooled human serum samples from 90 patients each with chronic obstructive pulmonary disease (COPD), lung cancer (LC), and their comorbidity (COPD with +LC) to healthy individuals (as control). Sample pooling was used to attenuate the influence of the individual samples on the resulting profiles. For better comparability and consequently easier structural identification, the timescale of the acquired electropherograms were converted from migration time to GU values using the GUcal software (freely available from GUcal.hu) [29]. A representative electropherogram of the pooled healthy human serum sample N- *N*-glycome is shown in Figure 1, featuring the separation of 61 peaks. Structural elucidation of all separated N- *N*-glycans utilized direct mining of GU database entries (GlycoStore.org), exoglycosidase digestion based carbohydrate sequencing [28], comparison to oligosaccharides released from carefully chosen glycoprotein standards (ribonuclease B, immunoglobulin G, αalpha-1-antitrypsin–), and some earlier published literature data on the same subject matter [30-32]. The exoglycosidase based glycan sequencing process is shown in Figure 2, utilizing sialidase A, galactosidase and hexosaminidase, depicted by the corresponding traces. Sequence information was derived from the GU value shifts of the individual peaks as the result of the consecutive exoglycosidase treatments [33].

Table 1 lists all identified N- *N*-glycan structures in the pooled human serum sample as numbered in Figure 1. The first level of structural elucidation of the separated glycans accomplished by using

their GU values to search the GlycoStore database (glycostore.org). Glycans denoted by # were identified considering the results of the sequential exoglycosidase digestion process shown in Figure 2. Structure identification of entries marked with asterisks were accomplished based on a comparative exercise utilizing the ~~N-~~ *N*-glycan profiles of IgG (*), RNase B (**) and AAT (***) [32, 34, 35]. It is important to note that besides these 61 carbohydrates, no other glycan structures were detected in the pathological samples compared to the control.

After the identification of the separated ~~N-~~ *N*-glycan structures in the pooled healthy human control serum sample, the peak areas of all peaks were computed and quantitatively compared to lung cancer, COPD and their comorbidity (COPD with ~~+~~LC) pooled sample data with their respective SD values based on the triplicate runs. Before the integration step for peak area determination, all electropherograms were normalized to Peak #61 (FA2BG2), i.e., the RFU values were divided by the RFU value of Peak #61, because it was apparently stable in size in all runs and well-separated from the other peaks. This normalization step was necessary to adequately calculate the peak areas of some of the not completely separated peaks. After the normalization step, the relative peak area % values were calculated by dividing the integrated peak areas by the sum of all integrated areas and multiplied by 100. Table 1 lists the suggested glycan structures corresponding to the separated peaks in Figure 1, the~~ir~~ calculated CE-LIF GU values and relative peak areas with their SD. Only peaks with >1% relative peak areas were taken into consideration in this comparative pilot study (highlighted bold in Table 1). All runs were done in triplicates and the average RSD of the relative peak areas was 3.46%. Please note that rows 62 - 78 in Table 1 represent the peaks appeared during exoglycosidase digestion after each sequential analysis step.

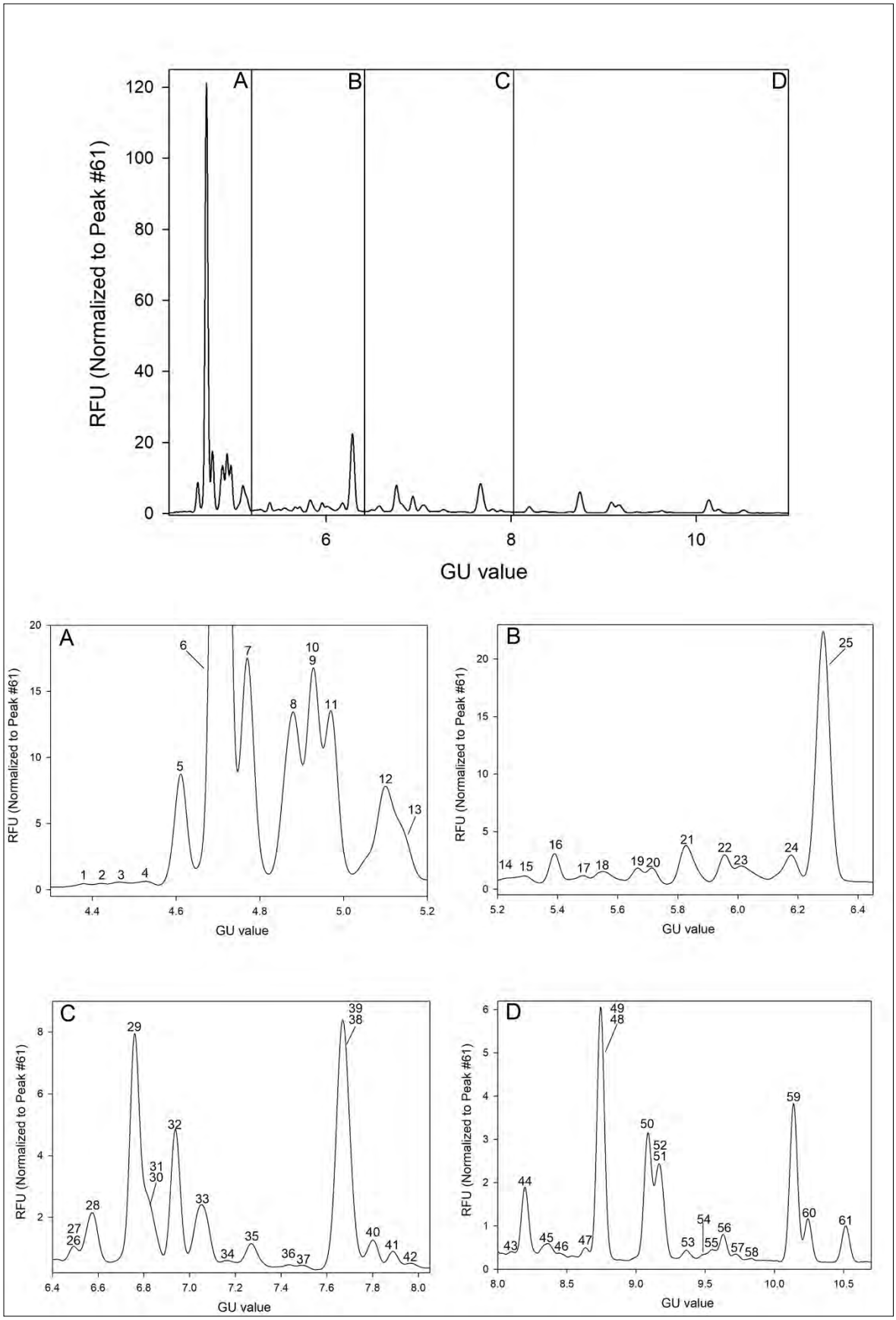

A
B
C
D
RFU (Normalized to Peak #61)
GU value
1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20 21 22 23 24 25
26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42
43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61

Figure 1. Capillary electrophoresis analysis of the released and APTS labeled ~~N-~~ *N*-glycans from pooled healthy human serum (upper panel). The lower panels show the enlarged sections of the electropherogram: A) GU 4.3-5.2; B) GU 5.2-6.4; C) GU 6.4-8.1; and D) GU 8.0-10.7. The corresponding structures, GU values and relative peak areas with their SD are listed in Table 1. Separation conditions: 40 cm effective capillary length (50 cm total length), 50 µm ID bare fused silica; 30 kV (0.17 min ramp time) separation voltage in reversed polarity mode. LIF detection (excitation: 488 nm/emission: 520 nm); separation temperature 30 °C. Injection: water pre-injection 5.0 s at 1.0 psi, followed by 3.0 kV/3.0 s sample and 2.0 kV/2.0 s bracketing standard (DP2 +DP15).

Table 1. The identified ~~N-~~ *N*-glycan structures in the pooled healthy control human serum sample, ~~with~~ their CE-LIF GU values and relative peak areas with SD. Bold entries represent peaks with >1% relative peak areas (at least in one sample groups from control, COPD, LC and, ~~LC+~~ with COPD). Entries from 62 to 78 depict the glycan structures of the sequencing analysis results (shown in Figure 2). Abbreviated structure names followed the nomenclature suggested in [36]. Entries were identified as follows: 1) based on their GU values using the GlycoStore database (glycostore.org); 2) considering the results of the sequential exoglycosidase digestion experiments shown in Figure 2 (denoted with #) and 3) all other glycan structures were identified based on *N*-glycosylation information of high abundant glycoproteins in serum: IgG (*), RNase B (**) and AAT (***) ~~N-glycosylation information~~.

	N-glycan structure	GU	Relative peak area (%)			N-glycan structure	GU	Relative peak area (%)	
1	A4G4[3,3,3,6]S4#	4.4	0.0955	±0.003	40	M6**	7.8	0.8338	±0.039
2	A4G4[3,3,3,3]S4#	4.46	0.0935	±0.002	41	A3G3[3]S1#	7.89	0.5296	±0.013
3	A1G1S1#	4.51	0.1307	±0.006	42	FA3#	7.98	0.4379	±0.011
4	FA4G4[3,3,3,6]S4***	4.57	0.1285	±0.004	43	A2B[6]G1	8.08	0.2962	±0.009
5	**FA4BG4[3,3,3,3]S4#**	**4.62**	**0.6814**	±0.015	**44**	**FA2B***	**8.2**	**1.9016**	±0.064
6	**A2G2[6]S2***	**4.71**	**21.8425**	±0.953	**45**	A2B[3]G1*	8.35	0.6615	±0.026
7	**FA3G3[6]S3#**	**4.77**	**0.9032**	±0.026	46	M7**	8.48	0.3190	±0.014
8	**A2G2[3]S2*****	**4.88**	**2.4776**	±0.098	47	M7**	8.62	0.3836	±0.013
9	**A2BG2S2***	**4.93**	**1.5508**	±0.038	**48**	**FA2[6]G1***	**8.75**	**6.1207**	±0.293
10	**M3**				**49**	**M7****			
11	**FA2G2S2***	**4.97**	**2.3546**	±0.052	**50**	**FA2[3]G1***	**9.09**	**2.9771**	±0.134
12	**FA2BG2S2***	**5.1**	**1.2156**	±0.026	**51**	**FA2B[6]G1***	**9.17**	**2.5049**	±0.067
13	**FA3G3[3]S3*****	**5.14**	**1.2645**	±0.031	**52**	**M8****			
14	A2[6]G1S1*	5.24	0.1604	±0.007	53	A4G4[6]S1	9.34	0.1668	±0.008
15	A2[3]G1[6]S1	5.3	0.1631	±0.005	54	A4G4[3]S1	9.48	0.1636	±0.004
16	FA3BG3S3#	5.39	0.5907	±0.026	55	FA2B[3]G1*	9.55	0.4411	±0.012
17	**A4G4[6]S3**	**5.47**	**0.8155**	±0.040	56	M8**	9.64	0.5983	±0.023
18	A2[3]G1[3]S1*	5.56	0.2984	±0.012	57	M8**	9.73	0.3001	±0.009

19	FA2G2[6]S2***	5.66	0.3043	±0.018	58	FA4***	9.86	0.0874	±0.004
20	**A3G3[6]S2**	**5.72**	**1.1683**	±0.031	**59**	**FA2G2***	**10.14**	**3.5100**	±0.089
21	FA2[6] G1S1*	5.83	0.7706	±0.029	60	M9**	10.25	0.9882	±0.042
22	**A4G4[3,3,3]S3**	**5.95**	**0.9457**	±0.024	61	FA2BG2*	10.52	0.8847	±0.035
23	FA2[3]G1S1*	6.03	0.5180	±0.017	62	A1G1	6.62	Sequencing	
24	**A3G3[3]S2**	**6.17**	**0.8196**	±0.030	63	A2[3]G1	8.12	Sequencing	
25	**A2G2[6]S1***	**6.28**	**13.1137**	±0.465	64	A2G2	9.17	Sequencing	
26	FA2B[6]G1S1*	6.48	0.4828	±0.016	65	A2BG2	10	Sequencing	
27	FA2B[3]G1S1*				66	A3[6]G3	11.36	Sequencing	
28	**A2BG2S1***	**6.58**	**1.3189**	±0.034	67	A3[3]G3	11.63	Sequencing	
29	**FA2G2S1***	**6.76**	**4.1359**	±0.095	68	FA3G3	12.33	Sequencing	
30	**A2***	**6.82**	**2.7863**	±0.119	69	FA3BG3	12.54	Sequencing	
31	**M5****				70	A4G4	13.72	Sequencing	
32	**FA2BG2S1***	**6.94**	**2.2966**	±0.063	71	FA4G4	14.38	Sequencing	
33	**A4G4[6]S2**	**7.05**	**1.3916**	±0.049	72	FA4BG4	14.45	Sequencing	
34	A4G4[3]S2	7.17	0.5179	±0.023	73	A1	5.86	Sequencing	
35	**A2B***	**7.27**	**0.7528**	±0.023	74	A3	7.81	Sequencing	
36	A2[6]G1	7.4	0.1440	±0.004	75	A4	8.94	Sequencing	
37	A3G3[6]S1	7.51	0.1470	±0.005	76	FA3B	9.93	Sequencing	
38	**FA2***	**7.67**	**9.5132**	±0.270	77	FA4B	11.02	Sequencing	
39	**M6****				78	FM3	5.09	Sequencing	

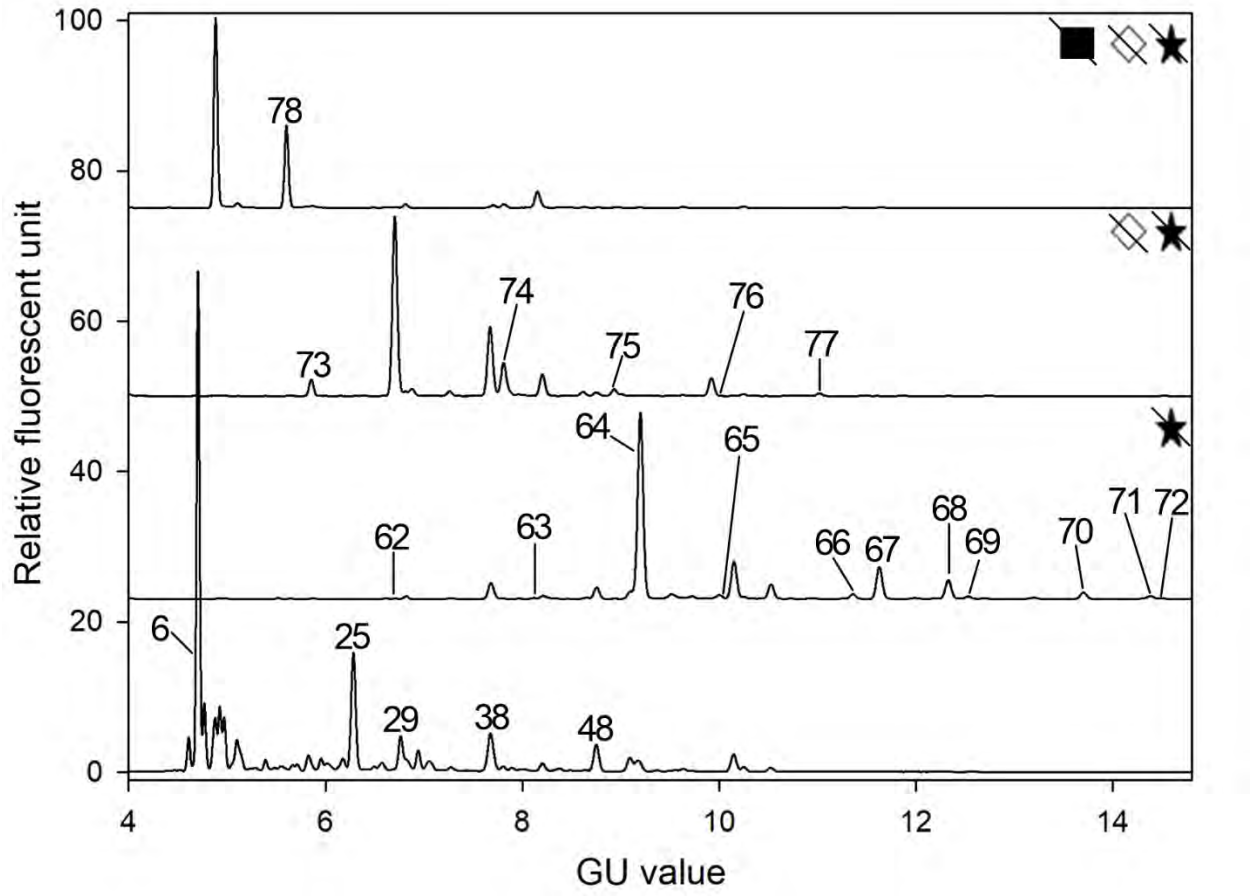


Figure 2. Exoglycosidase digestion based carbohydrate sequencing of the pooled healthy human control serum ~~N-~~ *N*-glycome using sialidase (★), galactosidase (◇), and hexosaminidase (■) to verify the identity of the glycan structures of interest. Separations conditions were the same as in Figure 1. The corresponding structures and GU values are listed in Table 1.

Changes in relative peak areas of the pooled lung cancer, COPD and comorbidity (COPD ~~+~~ with LC) samples were compared to the pooled healthy control, based on their capillary electrophoresis analysis results. As shown in Table 2, only glycans with significant alterations were taken into consideration during the evaluation process, i.e., ~~N-~~ *N*-glycan structures satisfying the following two criteria: 1) the relative peak areas of the ~~N-~~ *N*-linked glycan structures were >1%, at least in one sample group of healthy control, COPD, lung cancer, or their comorbidity (bold structures in

Table 1); and 2) at least one of the observed relative peak area differences between any of the disease groups and the control was >33%.

As one can observe, several relative peak area values in the pooled LC sample were considerably altered compared to the control (Table 2, Column 3), such as FA4BG4[3,3,3,3]S4 (peak 5: +167%±7.4), FA3G3[6]S3 (peak 7: +176%±9.7), A2BG2S2 + M3 (peaks 9-10: +133%±6.3), FA2G2S2 (peak 11: 105%±5.1), FA2BG2S2 (peak 12: 50%±2.9), A2BG2S1 (peak 28: -33%±1.6), FA2G2S1 (peak 29: -41%±2.1), FA2BG2S1 (peak 32: -41%±2.1), A2B (peak 35: +36%±2.1) and FA2G2 (peak 59: -36%±1.7). These changes may have diagnostic glycobiomarker potential for lung cancer. On the other hand, one should take into consideration that the following relative peak areas also increased in COPD: FA4BG4[3,3,3,3]S4 (peak 5, LC:+167%±7.4, COPD: +106%±5.2), A2BG2S2 + M3 (peaks 9-10, LC:+133%±6.3, COPD: +58%±2.8) and FA2BG2S2 (peak 12, LC:+50%±2.9, COPD: +33%±1.8), therefore, they may indicate the presence of both lung cancer or COPD, but not the comorbidity. In addition the increment of the relative peak area of FA3G3[6]S3 (peak 7, LC:+176%±9.7, COPD: +79%±3.8, LC+COPD: +56%±3.0) could represent either lung cancer or COPD or their comorbidity. ~~Furthermore~~Interestingly, peak 28, corresponding to the A2BG2S1 structure slightly increased (+12%±0.6) in comparison of COPD to control, while significantly decreased (-33%±1.6) in lung cancer compared to control, therefore could also be a glycobiomarker for lung cancer. Moreover, ~~it could also help~~ using Column 4 in Table 2, one can ~~to~~ distinguish COPD from lung cancer considering the increasing changes in FA3G3[6]S3, A2BG2S2+ M3, FA2G2S2, A2B and FA2 + M6 (54%±2.5, 47%±2.2, 79%±4.2, 35%±2.0 and 41%±2.1), and the decreasing values of A2BG2S1 and FA2G2S1 (41%±1.9 and 39%±2.0).~~.~~

In addition, the relative peak area change of peak 29 (FA2G2S1) may discern comorbidity of COPD with lung cancer (-37 %±1.7) from COPD (-4%±0.2), but not from lung cancer (-41%±2.1). In all pathological samples peak 59 (FA2G2, LC: -36%±1.7, COPD: -34%±1.6, LC+COPD: -39%±2.1) decreased, thus, possibly~~e can~~ distinguish them only from the control. Please note that FA2 + M6 (peaks 38-39) and A2BG2S2 + M3 (peaks 9-10) structures were co-migrating, thus their individual alterations could not be precisely evaluated.

Table 2. Comparison of the relative peak areas of lung cancer (LC), COPD, ~~and~~ their comorbidity (COPD~~+~~ with LC) to the control sample and between LC and COPD, where at least one observed difference was >33% (bold numbers) for any peak with >1% relative area between the disease and control samples.

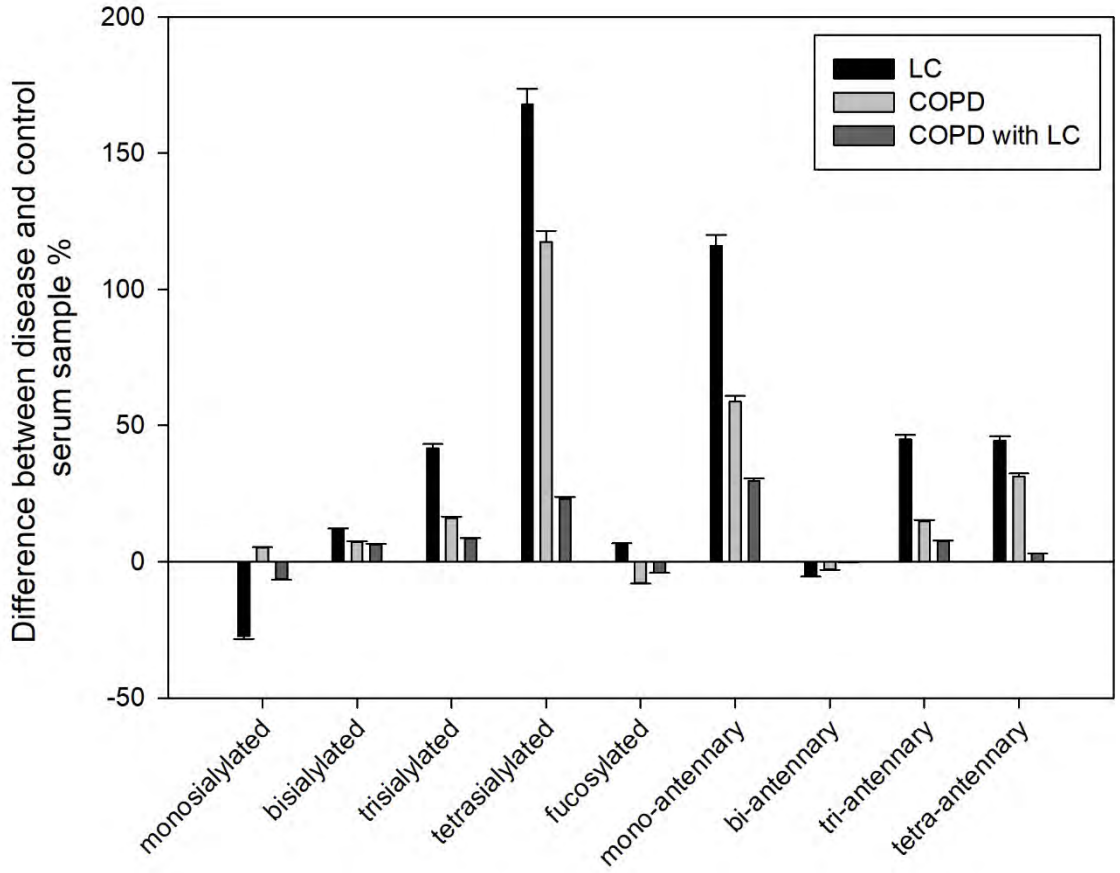
Peak Number	N-Glycan structure	LC to Control %	COPD to Control %	LC + COPD to Control %	LC to COPD %
5	FA4BG4[3,3,3,3]S4	**167 ±7.4**	**106 ±5.2**	18 ±0.9	29 ±1.4
7	FA3G3[6]S3	**176 ±9.7**	**79 ±3.8**	**56 ±3.0**	**54 ±2.5**
9-10	A2BG2S2, M3	**133 ±6.3**	**58 ±2.8**	29 ±1.4	**47 ±2.2**
11	FA2G2S2	**105 ±5.1**	15 ±0.7	4 ±0.2	**79 ±4.2**
12	FA2BG2S2	**50 ±2.9**	**33 ±1.8**	16 ±0.7	13 ±0.9
28	A2BG2S1	**-33 ±1.6**	12 ±0.6	-17 ±1.0	**-41 ±1.9**
29	FA2G2S1	**-41 ±2.1**	-4 ±0.2	**-37 ±1.7**	**-39 ±2.0**
32	FA2BG2S1	**-41 ±2.1**	-30 ±1.5	-26 ±1.3	-17 ±0.8
35	A2B	**36 ±2.1**	0 ±0	12 ±0.6	**35 ±2.0**
38-39	FA2, M6	15 ±0.9	-18 ±0.9	28 ±1.6	**41 ±2.1**
59	FA2G2	**-36 ±1.7**	**-34 ±1.6**	**-39 ±2.1**	-3 ±0.1

Figure 3 shows the quantitative alterations among the relative peak areas within the ~~N~~ *N*-glycan subclasses specified by branching, degrees of sialylation and total fucosylation, determined in the control, COPD, lung cancer and comorbidity of COPD with lung cancer patient serum samples.

Variations in the size of the carbohydrate structures were evaluated based on their degree of branching, i.e., the amounts of various mono-, bi-, tri-, and tetra-antennary types. The number of terminal sialic acids was also taken into consideration in the evaluation process, focusing on the tendency of the relative peak area changes in mono-, bi-, tri- and tetra-sialylation. The relative peak areas of mono-sialylated glycans slightly increased in COPD (+5%±0.2), slightly decreased in comorbidity of COPD with lung cancer (-6%±0.2) and more substantially decreased in lung cancer (-27%±0.9). Moreover, the relative peak areas of bi-, tri- and tetrasialylated subclasses variably increased in COPD, lung cancer and their comorbidity (COPD + LC) in comparison to the pooled control healthy human serum and the increment was the highest in lung cancer (bi-antennary: +12%±0.4, tri-antennary: +42%±1.4, tetra-antennary: +168%±5.8), medium in COPD (bi-antennary: +7%±0.2, tri-antennary: +16%±0.6, tetra-antennary: +117%±4.1) and the smallest in their comorbidity (bi-antennary: +6%±0.2, tri-antennary: +8%±0.3, tetra-antennary: +23%±0.8). The relative peak areas of fucosylated ~~N-~~ *N*-glycans slightly decreased in COPD (-8%±0.3) and also in comorbidity of COPD with lung cancer (-4%±0.1) in contrast to lung cancer, where a small increase (+7%±0.2) was observed.

The relative peak areas of mono-, tri- and tetra-antennary glycans similarly to the bi-, tri- and tetra-sialylated subclasses increased in all three disease groups compared to the healthy control with the highest increment in lung cancer (mono-antennary: +116%±4.0, tri-antennary: +45%±1.6, tetra-antennary: 44.5%±1.5), medium in COPD (mono-antennary: +59%±2.0, tri-antennary: +15%±0.5, tetra-antennary: 31%±1.1) and the smallest in their comorbidity (mono-antennary: +30%±1.0, tri-antennary: +7.5%±0.3, tetra-antennary: 3%±0.1). In contrary, the relative peak areas of bi-antennary ~~N-~~ *N*-glycans hardly decreased or not changed at all in all sample types of COPD (-

3%±0.1), lung cancer (-5%±0.2) and comorbidity of COPD with lung cancer (-0.3%±0.0), compared to the control. However, taking a closer look at the data, these observed changes were apparently caused by the interplay between various increasing and decreasing amounts of bi-antennary structures, i.e., the relative peak areas of some bi-antennary glycans significantly decreased (e.g., A2BG2S1 (peak 28): LC: -33%±1.6, LC+COPD: -17%±1.0; FA2G2S1 (peak 29): LC: -41%±2.1, COPD: -4%±0.2, LC+COPD: -37%±1.7; FA2G2 (peak 59): LC: -36%±1.7, COPD: -34%±1.6, LC+COPD: -39%±2.1); while others increased (e.g., A2BG2S2 + M3) (peaks 9-10): LC: +133%±6.3, COPD: +58%±2.8, LC+COPD: +29%±1.4; FA2G2S2 (peak 11): LC: +105%±5.1, COPD: +15%±0.7, LC+COPD: +4%±0.2; FA2BG2S2 (peak 12): LC: +50%±2.9, COPD: +33%±1.8, LC+COPD: +16%±0.7 and A2B (peak 35): LC: +36%±2.1, LC+COPD: +12%±0.6), thus their changes were almost counterbalanced. Based on these observations, as a first approximation we consider that ~~N-~~ *N*-glycan subclasses of mono- tri- and tetrasialylated, fucosylated as well as mono-, tri- and tetra-antennary forms could serve as potential biomarkers for lung cancer, COPD and their comorbidity.


Difference between disease and control serum sample %
200
150
100
50
0
-50
LC
COPD
COPD with LC
monosialylated
bisialylated
trisialylated
tetrasialylated
fucosylated
mono-antennary
bi-antennary
tri-antennary
tetra-antennary

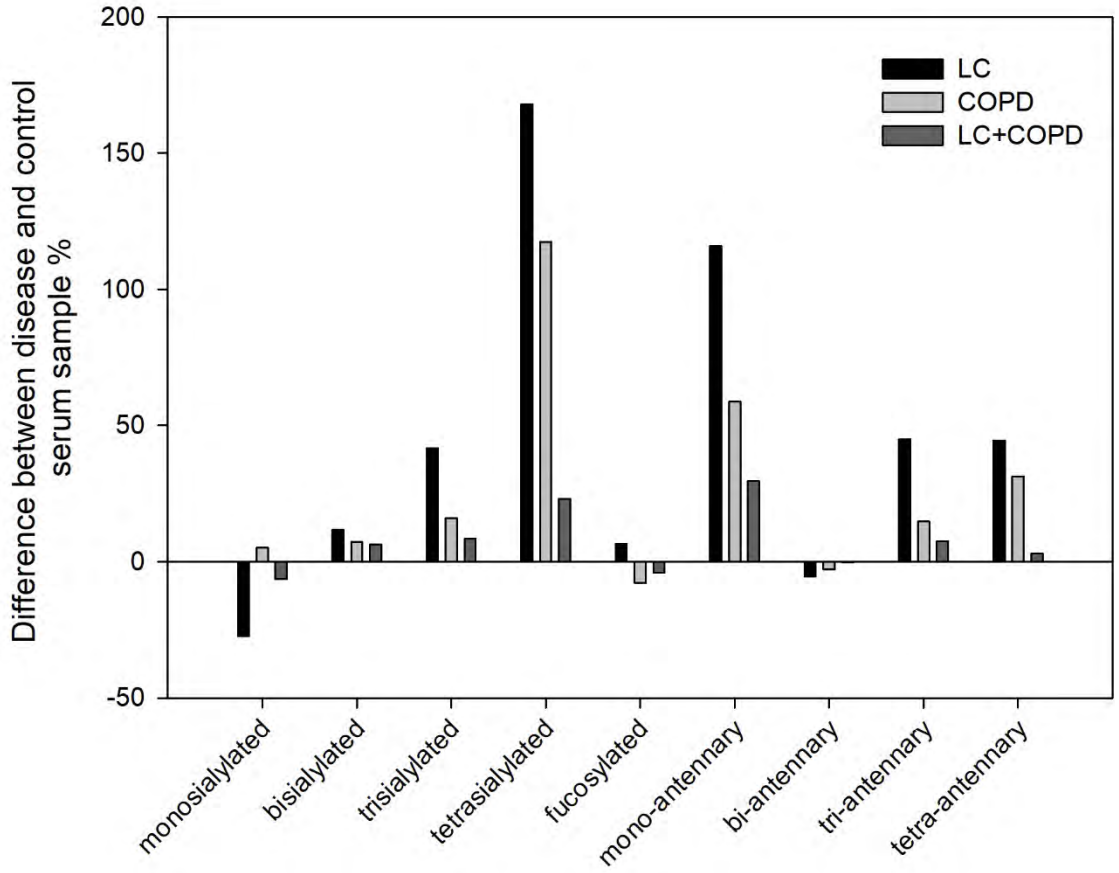


Figure 3. Alterations among the relative peak areas of specific ~~N-~~ *N*-glycan subclasses (Sialoforms: mono-, bi-, tri- and tetra-sialo; Fucosylated and Branching: mono-, bi- tri- and tetra-antennary) of lung cancer (black), COPD (gray) and their comorbidity (dark gray) with their corresponding RSDs. The results were calculated based on the data in supplementary Table 1.

Conclusion

In this ~~work~~pilot study, the ~~N-~~ *N*-glycosylation profiles of patient samples of chronic inflammatory (COPD) and malignant (LC) pulmonary diseases as well as their comorbidity (COPD ~~+~~with LC) were quantitatively studied and compared to healthy controls. A novel temperature adjusted denaturation protocol and extended enzymatic release and evaporative derivatization time was used for the ~~The~~ asparagine linked oligosaccharides from the complex serum samples, which were then analyzed by capillary electrophoresis with high sensitivity laser induced fluorescence

detection. Sixty-one ~~N-~~ *N*-glycan structures were identified in the control human serum samples and since no other glycans appeared in any of the three disease categories, only these 61 structures were quantitatively monitored in this study. Our results suggested that certain serum ~~N-~~ *N*-glycans could be used as potential markers for the different types of pulmonary diseases. Therefore, the panel of the 13 glycans listed in Table 2 could be considered to differentiate lung cancer, COPD and their comorbidity from the control and LC from COPD. In addition, alterations in the ~~N-~~ *N*-glycan subclasses, such as fucosylated, mono-, bi-, tri- and tetra-sialylo, as well as mono-, bi-, tri- and tetra-antennary glycans could also carry interesting diagnostic information. The glycan panel in Table 2 and the corresponding subclasses ~~information~~ may provide even more reliable information as they represent the sum of multiple structural changes caused by a given disease. This is especially applicable for the highly branched sialylated structures as our recent genotyping data suggested significant increase of MGAT5 activity, i.e., increased branching in lung cancer [37]. Currently we are in the process of collecting 300 samples from each the disease groups that will be individually analyzed in view of our preliminary pooled sample based results.

Acknowledgement

The authors gratefully acknowledge the support of the National Research, Development and Innovation Office (NKFIH) (K 116263) grants of the Hungarian Government. This work was also supported by the BIONANO_GINOP-2.3.2-15-2016-00017 project, the V4-Korea Joint Research Program, project National Research, Development and Innovation Office (NKFIH) (NN 127062) grants of the Hungarian Government and the Human Capacities Grant Management Office, the ÚNKP-18-3-I-DE-393 New National Excellence Program Hungarian Ministry of Human Capacities and the János Bolyai Research Scholarship of the Hungarian

Academy of Sciences. This is contribution #152 from the Horvath Csaba Memorial Laboratory of Bioseparation Sciences.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin., 68 (2018) 394-424.

[2] A.L. Durham, I.M. Adcock, The relationship between COPD and lung cancer, Lung Cancer, 90 (2015) 121-127.

[3] J.L. López-Campos, W. Tan, J.B. Soriano, Global burden of COPD, Respirology, 21 (2016) 14-23.

[4] D. Spyratos, E. Papadaki, S. Lampaki, T. Kontakiotis, Chronic obstructive pulmonary disease in patients with lung cancer: prevalence, impact and management challenges, Lung Cancer (Auckl), 8 (2017) 101-107.

[5] Y. Wang, F. Jiang, X. Tan, P. Tian, CT-guided percutaneous transthoracic needle biopsy for paramediastinal and nonparamediastinal lung lesions: Diagnostic yield and complications in 1484 patients, Medicine, 95 (2016) e4460-e4460.

[6] K. Shyamala, H.C. Girish, S. Murgod, Risk of tumor cell seeding through biopsy and aspiration cytology, J. Int. Soc. Prev. Community Dent, 4 (2014) 5-11.

[7] Y. Xi, J. Fan, D. Che, K. Zhai, T. Ren, X. Feng, L. Shang, J. Hu, Y. Yu, Q. Meng, Distant Metastasis and Survival Outcomes after Computed Tomography-Guided Needle Biopsy in Resected Stage I-III Non-Small Cell Lung Cancer, J. Cancer, 8 (2017) 3356-3361.

[8] O. Gornik, T. Pavic, G. Lauc, Alternative glycosylation modulates function of IgG and other proteins - implications on evolution and disease, Biochim. Biophys. Acta, 1820 (2012) 1318-1326.

[9] T. Suzuki, K. Ohtsubo, N. Taniguchi, Sugar Chains: Decoding the Functions of Glycans, 2015.

[10] H. Lemjabbar-Alaoui, A. McKinney, Y.-W. Yang, V.M. Tran, J.J. Phillips, Chapter Nine - Glycosylation Alterations in Lung and Brain Cancer, in: R.R. Drake, L.E. Ball (Eds.) Advances in Cancer Research, Academic Press 2015, pp. 305-344.

[11] F. Clerc, K.R. Reiding, B.C. Jansen, G.S. Kammeijer, A. Bondt, M. Wuhrer, Human plasma protein N-glycosylation, Glycoconj. J., 33 (2016) 309-343.

[12] A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, H.P. Packer, J.H. Prestegard, R.L. Schnaar, P.H. Seeberger, Essentials of Glycobiology, Cold Spring Harbor Laboratory Press, DOI (2015-2017).
[13] R.R. Townsend, A.T. Hotchkiss, Techniques in Glycobiology, Marcel Dekker Inc., New York, NY, 1997.
[14] R. Saldova, L. Royle, C.M. Radcliffe, U.M. Abd Hamid, R. Evans, J.N. Arnold, R.E. Banks, R. Hutson, D.J. Harvey, R. Antrobus, S.M. Petrescu, R.A. Dwek, P.M. Rudd, Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG, Glycobiology, 17 (2007) 1344-1356.
[15] T. Pavic, D. Dilber, D. Kifer, N. Selak, T. Keser, D. Ljubicic, A. Vukic Dugac, G. Lauc, L. Rumora, O. Gornik, N-glycosylation patterns of plasma proteins and immunoglobulin G in chronic obstructive pulmonary disease, J. Transl. Med., 16 (2018) 323.
[16] A. Bondt, S. Nicolardi, B.C. Jansen, K. Stavenhagen, D. Blank, G.S. Kammeijer, R.P. Kozak, D.L. Fernandes, P.J. Hensbergen, J.M. Hazes, Y.E. van der Burgt, R.J. Dolhain, M. Wuhrer, Longitudinal monitoring of immunoglobulin A glycosylation during pregnancy by simultaneous MALDI-FTICR-MS analysis of N- and O-glycopeptides, Sci. Rep., 6 (2016) 27955.
[17] A. Bondt, S. Nicolardi, B.C. Jansen, T.M. Kuijper, J.M.W. Hazes, Y.E.M. van der Burgt, M. Wuhrer, R.J.E.M. Dolhain, IgA N- and O-glycosylation profiling reveals no association with the pregnancy-related improvement in rheumatoid arthritis, Arthritis research & therapy, 19 (2017) 160-160.
[18] Y. Liang, T. Ma, A. Thakur, H. Yu, L. Gao, P. Shi, X. Li, H. Ren, L. Jia, S. Zhang, Z. Li, M. Chen, Differentially expressed glycosylated patterns of alpha-1-antitrypsin as serum biomarkers for the diagnosis of lung cancer, Glycobiology, 25 (2015) 331-340.
[19] T. Fujimura, Y. Shinohara, B. Tissot, P.C. Pang, M. Kurogochi, S. Saito, Y. Arai, M. Sadilek, K. Murayama, A. Dell, S. Nishimura, S.I. Hakomori, Glycosylation status of haptoglobin in sera of patients with prostate cancer vs. benign prostate disease or normal subjects, Int. J. Cancer, 122 (2008) 39-49.
[20] C. Varadi, S. Mittermayr, A. Szekrenyes, J. Kadas, L. Takacs, I. Kurucz, A. Guttman, Analysis of haptoglobin N-glycome alterations in inflammatory and malignant lung diseases by capillary electrophoresis, Electrophoresis, 34 (2013) 2287-2294.

[21] O. Choi, N. Tomiya, J.H. Kim, J.M. Slavicek, M.J. Betenbaugh, Y.C. Lee, N-glycan structures of human transferrin produced by Lymantria dispar (gypsy moth) cells using the LdMNPV expression system, Glycobiology, 13 (2003) 539-548.
[22] A. Quaranta, A. Sroka-Bartnicka, E. Tengstrand, G. Thorsén, N-Glycan profile analysis of transferrin using a microfluidic compact disc and MALDI-MS, Analytical and bioanalytical chemistry, 408 (2016) 4765-4776.
[23] E. Ito, R. Oka, T. Ishii, H. Korekane, A. Kurimoto, Y. Kizuka, S. Kitazume, S. Ariki, M. Takahashi, Y. Kuroki, K. Kida, N. Taniguchi, Fucosylated surfactant protein-D is a biomarker candidate for the development of chronic obstructive pulmonary disease, J. Proteomics, 127 (2015) 386-394.
[24] J.N. Arnold, R. Saldova, M.C. Galligan, T.B. Murphy, Y. Mimura-Kimura, J.E. Telford, A.K. Godwin, P.M. Rudd, Novel Glycan Biomarkers for the Detection of Lung Cancer, J. Proteome Res., 10 (2011) 1755-1764.
[25] L.R. Ruhaak, U.T. Nguyen, C. Stroble, S.L. Taylor, A. Taguchi, S.M. Hanash, C.B. Lebrilla, K. Kim, S. Miyamoto, Enrichment strategies in glycomics-based lung cancer biomarker development, Proteomics Clin. Appl., 7 (2013) 664-676.
[26] E.F. Schisterman, A. Vexler, To pool or not to pool, from whether to when: applications of pooling to biospecimens subject to a limit of detection, Paediatr Perinat Epidemiol, 22 (2008) 486-496.
[27] B. Reider, M. Szigeti, A. Guttman, Evaporative fluorophore labeling of carbohydrates via reductive amination, Talanta, 185 (2018) 365-369.
[28] A. Guttman, K.W. Ulfelder, Exoglycosidase matrix-mediated sequencing of a complex glycan pool by capillary electrophoresis, J. Chromatogr. A, 781 (1997) 547-554.
[29] G. Jarvas, M. Szigeti, A. Guttman, GUcal: An integrated application for capillary electrophoresis based glycan analysis, Electrophoresis, 36 (2015) 3094-3096.
[30] R. Saldova, A. Asadi Shehni, V.D. Haakensen, I. Steinfeld, M. Hilliard, I. Kifer, A. Helland, Z. Yakhini, A.L. Borresen-Dale, P.M. Rudd, Association of N-glycosylation with breast carcinoma and systemic features using high-resolution quantitative UPLC, J. Proteome Res., 13 (2014) 2314-2327.
[31] L. Royle, M.P. Campbell, C.M. Radcliffe, D.M. White, D.J. Harvey, J.L. Abrahams, Y.-G. Kim, G.W. Henry, N.A. Shadick, M.E. Weinblatt, D.M. Lee, P.M. Rudd, R.A. Dwek, HPLC-based

analysis of serum N-glycans on a 96-well plate platform with dedicated database software, Analytical Biochemistry, 376 (2008) 1-12.
[32] K.J. Lee, S.M. Lee, J.Y. Gil, O. Kwon, J.Y. Kim, S.J. Park, H.S. Chung, D.B. Oh, N-glycan analysis of human alpha1-antitrypsin produced in Chinese hamster ovary cells, Glycoconj. J., 30 (2013) 537-547.
[33] M. Szigeti, A. Guttman, Automated N-Glycosylation Sequencing Of Biopharmaceuticals By Capillary Electrophoresis, Sci. Rep., 7 (2017) 11663.
[34] A. Guttman, S. Herrick, Effect of the quantity and linkage position of mannose (alpha 1,2) residues in capillary gel electrophoresis of high-mannose-type oligosaccharides, Anal. Biochem., 235 (1996) 236-239.
[35] S. Mittermayr, J. Bones, A. Guttman, Unraveling the glyco-puzzle: glycan structure identification by capillary electrophoresis, Anal. Chem., 85 (2013) 4228-4238.
[36] D.J. Harvey, A.H. Merry, L. Royle, M.P. Campbell, P.M. Rudd, Symbol nomenclature for representing glycan structures: Extension to cover different carbohydrate types, Proteomics, 11 (2011) 4291-4295.
[37] Z. Elek, Z. Kovács, G. Keszler, M. Szabo, E. Csanky, J. Luo, A. Guttman, Z. Rónai, High throughput multiplex SNP-analysis in chronic obstructive pulmonary disease and lung cancer, Current Molecular Mecicine, 19 (2019) In Press.

Highlights:

- Comparative analysis of the human serum N-Glycome between lung cancer, COPD and their comorbidity to control.
- Sixty-one N-glycan structures were identified in the pooled control human serum sample and compared to pooled LC, COPD and COPD+LC patient samples.
- A panel of 13 N-glycan structures were identified as potential glycobiomarkers to reveal significant changes between the pathological and control samples.
- N-glycan profile changes of specific N-glycan subclasses of branching, fucosylation, etc., may also hold important glycobiomarker value.

Comparative analysis of the human serum N-glycome in lung cancer, COPD and their comorbidity using capillary electrophoresis

Brigitta Mészáros[1], Gábor Járvás[1,2], Anna Farkas[1], Márton Szigeti[1,2], Zsuzsanna Kovács[1], Renáta Kun[1,3], Miklós Szabó[3], Eszter Csánky[3] and András Guttman[1,2]

[1]Horvath Csaba Memorial Laboratory of Bioseparation Sciences, Research Center for Molecular Medicine, Doctoral School of Molecular Medicine, Faculty of Medicine, University of Debrecen, Debrecen, Hungary

[2]Research Institute of Biomolecular and Chemical Engineering, University of Pannonia, Veszprem, Hungary

[3]Department of Pulmonology, Semmelweis Hospital, Miskolc, Hungary

Abstract

Lung cancer (LC) and chronic obstructive pulmonary disease (COPD) are prevalent ailments with a great challenge to distinguish them based on symptoms only. Since they require different treatments, it is important to find non-invasive methods capable to readily diagnose them. Moreover, COPD increases the risk of lung cancer development, leading to their comorbidity. In this study the N-glycosylation profile of pooled human serum samples (90 patients each) from lung cancer, COPD and comorbidity (LC+COPD) patients were investigated in comparison to healthy individuals (control) by capillary gel electrophoresis with high sensitivity laser-induced fluorescence detection. Sixty-one N-glycan structures were identified in the pooled control human serum sample and the profile was compared to pooled lung cancer, COPD and comorbidity of COPD with lung cancer patient samples. It is important to note that no other sugar structures were detected in any of the patient groups, only quantitative differences were observed. Based on this comparative exercise, a panel of 13 N-glycan structures were identified as potential glycobiomarkers to reveal significant changes (>33% in relative peak areas) between the pathological and control samples. In addition to N-glycan profile changes, alterations in the

individual N-glycan subclasses, such as total fucosylation, degree of sialylation and branching may also hold important glycobiomarker value.

Comparative analysis of the human serum *N*-glycome in lung cancer, COPD and their comorbidity using capillary electrophoresis

Brigitta Mészáros[1], Gábor Járvás[1,2], Anna Farkas[1], Márton Szigeti[1,2], Zsuzsanna Kovács[1], Renáta Kun[1,3], Miklós Szabó[3], Eszter Csánky[3] and András Guttman[1,2]

[1]Horváth Csaba Memorial Laboratory of Bioseparation Sciences, Research Center for Molecular Medicine, Doctoral School of Molecular Medicine, Faculty of Medicine, University of Debrecen, Debrecen, Hungary

[2]Research Institute of Biomolecular and Chemical Engineering, University of Pannonia, Veszprem, Hungary

[3]Department of Pulmonology, Semmelweis Hospital, Miskolc, Hungary

Keywords: human serum, *N*-linked glycan, capillary electrophoresis, pulmonary diseases

Abbreviations:

COPD – chronic obstructive pulmonary disease
LC – lung cancer
DP – degree of polymerization
APTS – 8-aminopyrene-1,3,6-trisulfonic acid
GU – glucose unit
IgG – immunoglobulin G
IgA – immunoglobulin A
AAT – α-1-antitrypsin
HP – haptoglobin
TR – transferrin
DTT – dithiothreitol
RNase B – ribonuclease B
RFU –relative fluorescence unit

Abstract

Lung cancer (LC) and chronic obstructive pulmonary disease (COPD) are prevalent ailments with a great challenge to distinguish them based on symptoms only. Since they require different treatments, it is important to find non-invasive methods capable to readily diagnose them. Moreover, COPD increases the risk of lung cancer development, leading to their comorbidity. In this pilot study the *N*-glycosylation profile of pooled human serum samples (90 patients each) from

lung cancer, COPD and comorbidity (LC with COPD) patients were investigated in comparison to healthy individuals (control) by capillary gel electrophoresis with high sensitivity laser-induced fluorescence detection. Sample preparation was optimized for human serum samples introducing a new temperature adjusted denaturation protocol to prevent precipitation and increased endoglycosidase digestion time to assure complete removal of the *N*-linked carbohydrates. Sixty-one *N*-glycan structures were identified in the pooled control human serum sample and the profile was compared to pooled lung cancer, COPD and comorbidity of COPD with lung cancer patient samples. One important findings was that no other sugar structures were detected in any of the patient groups, only quantitative differences were observed. Based on this comparative exercise, a panel of 13 *N*-glycan structures were identified as potential glycobiomarkers to reveal significant changes (>33% in relative peak areas) between the pathological and control samples. In addition to *N*-glycan profile changes, alterations in the individual *N*-glycan subclasses, such as total fucosylation, degree of sialylation and branching may also hold important glycobiomarker values.

Introduction

Lung cancer (LC) is one of the predominantly diagnosed cancers (11.6% of the total cases) and was the main cause of cancer deaths (18.4% of the total cancer deaths) in 2018 in both genders combined [1]. Another prevalent lung disease is Chronic Obstructive Pulmonary Disease (COPD), mostly caused by cigarette smoking, representing a great but certainly preventable risk factor [2]. COPD has recently become in the spotlight due to its elevating incidence rate, morbidity, mortality and increasing the risk of developing lung cancer, in many cases creating formidable challenges for the global healthcare systems [3]. Simultaneous diagnoses of LC and COPD (i.e., comorbidity) represent poor prognosis [2], thus treatments should be planned accordingly [4]. The high mortality rate of COPD and lung cancer could be due to the asymptomatic property at their early stages, and

the lack of appropriate distinguishing diagnostic tools. While biopsy can differentiate malignant (LC) and benign (inflammation) lesions that imaging techniques are not capable in a reliable diagnostic manner, it is an invasive process with a number of risk factors including infections, pneumonia, pneumothorax, bleeding and haemoptysis, just to list the most frequent ones [5]. Furthermore, another potential drawback of biopsy is the risk of tumor spreading, i.e., during the biopsy process, possible metastasis initiating circulating tumor cells may get away from the primary tumor [6, 7]. Therefore, there is an urgent need to develop non-invasive molecular diagnostic tools capable of predicting the presence and prognosis of the actual disease (LC, COPD or comorbidity) with adequate specificity and sensitivity.

Ever-growing evidence shows the importance of the *N*-glycosylation of proteins in biological systems, demonstrating that this post-translational modification is as essential as the polypeptide backbone itself, playing significant roles in forming their higher order structure, biochemical properties and function [8]. Suzuki et al. [9] studied the importance and roles of *N*-glycosylation in COPD and found that the reduction of FUT8 activity has close relation with the progression of the disease. Phillips and coworkers [10] reviewed the glycosylation aspects of lung cancer and the mechanisms of post-synthetic glycan modification during malignant transformation suggesting promising biomarkers and therapeutic possibilities based on their *N*-glycosylation alterations. The significance of these changes revealed that identifying *N*-glycan biomarkers as potential early detection markers and monitoring the treatment of lung diseases [2] can be of high importance. The human serum contains a plethora of proteins with the majority of them glycosylated [11]. Protein glycosylation is the consequence of very complex biochemical processes, regulated by a number of glycosidases and glycosyltransferases [9] resulting in diverse but protein specific glycan

profiles, which affect several cellular properties such as signaling, adhesion, motility and half-life, just to mention a few important ones [12]. Genetic and environmental factors can both affect the activity and level of this glycosylation machinery that may lead to altered glycan structures, possibly specific for pathological changes. Comprehensive analysis of complex carbohydrates requires high sensitivity methods with enhanced resolution because of their great structural diversity. The most frequently applied techniques for *N*-glycan analysis are high-performance liquid chromatography (HPLC), capillary electrophoresis (CE), mass spectrometry (MS) and NMR, or the combination of those such as LC-MS or CE-MS [13].

In several cases, *N*-glycosylation alterations affect the high abundant acute phase proteins, including immunoglobulin G (IgG) [14, 15], immunoglobulin A (IgA) [11, 16, 17], alpha-1-antitrypsin (AAT) [18], haptoglobin (HP) [19, 20] and transferrin (TR) [21, 22]. Pavić et al [15] analyzed the human plasma and its IgG subset of COPD patients and healthy controls by UPLC and provided new insights into plasma protein and IgG *N*-glycome changes. The plasma protein *N*-glycome significantly decreased in low branched type and increased in more complex glycan structures in COPD patients. Ito et al. investigated the *N*-glycans of a lung-specific protein, surfactant protein D, using matrix-assisted laser desorption/ionization quadrupole ion trap time-of-flight mass spectrometry and found that the fucosylation level was greatly elevated in COPD patients compared to controls [23]. Rudd and coworkers analyzed the serum of lung cancer patients and healthy controls by high performance hydrophilic interaction liquid chromatography (HILIC) and anion exchange chromatography [24]. They found that the level of bi- (A2), tri- (A3) and tetra-antennary (A4) glycans were significantly increased in lung cancer compared to healthy controls and a reduction was observed in the amounts of core-fucosylated bi-antennary structures. In

addition to total human serum, one of the abundant proteins, haptoglobin was analyzed and similar alterations were found in both sample types. Váradi et al. also studied the human serum haptoglobin in lung diseases and emphasized the importance of measuring the core- to arm-fucosylation ratios [20]. The slight decrease in the total fucosylation level of haptoglobin in the serum of COPD and pneumonia patients compared to the control group was the result of a significant decrease in arm-fucosylation and a slight increase in core-fucosylation. Elevated amounts of core- (FA4G4) and antennary-fucosylated tetra-antennary glycans (A4FG4) in haptoglobin were also observed comparing lung cancer and COPD patient groups. Ruhaak et al. investigated the *N*-glycans of several high abundance glycoproteins enriched by affinity capture from plasma samples of lung adenocarcinoma patients as well as healthy controls and analyzed with nano-HPLC-chip-TOF-MS to search for lung cancer biomarkers [25]. They found that, while the *N*-glycan profiles of blood-derived glycoproteins could be adequate biomarkers for lung cancer, protein enrichment did not improve specificity, but made the method more complicated. Liang et al. [18] examined different types of pulmonary diseases including lung adenocarcinoma, squamous cell lung cancer, small-cell lung cancer as well as several benign ones, such as pneumonia, pulmonary nodule and tuberculous pleuritis and suggested that the *N*-glycosylation patterns of α-1-antitrypsin could be a potential lung cancer biomarker.

In this paper we report on a pilot study of the *N*-glycosylation profiles of pooled human serum samples from patients with COPD, lung cancer and their comorbidities (COPD with LC) compared to control healthy subjects (pool of 90 patients each) using capillary electrophoresis with laser-induced fluorescent detection (CE-LIF). A new temperature adjusted denaturation protocol was used to prevent precipitation and the endoglycosidase digestion time was increased to assure full

removal of the *N*-linked oligosaccharides. Sixty-one *N*-glycans were identified in the pooled, control human serum samples based on data in publicly available GU databases and exoglycosidase digestion based carbohydrate sequencing. The relative peak areas of the identified *N*-glycans were used in a comparative quantitative evaluation to find a potential preliminary glycobiomarker panel to differentiate lung cancer, COPD and their comorbidity from each other and from the control.

Materials and methods

Chemicals and reagents

Water (HPLC grade), acetonitrile, sodium cyanoborohydride (1 M in THF), glycerol, ribonuclease B, immunoglobulin G, α-1-antitrypsin and DTT (dithiothreitol) were obtained from Sigma Aldrich (St. Louis, MO, USA). SDS (sodium dodecyl sulfate) and Nonidet P-40 were from VWR (Radnor, PA, USA). The Fast Glycan Labeling and Analysis Kit was from SCIEX (Brea, CA, USA) including the bracketing standards of maltose (DP2) and maltopentadecaose (DP15). The exoglycosidase enzymes of Sialidase A (*Arthrobacter ureafaciens*), β-Galactosidase (*Jack bean*) and β-N-Acetyl Hexosaminidase (*Jack bean*) were from ProZyme (Hayward, CA, USA). The endoglycosidase PNGase F was from Asparia Glycomics, (San Sebastian, Spain).

Sample preparation

All serum samples were collected with the appropriate Ethical Permissions (approval number: 23580-1/2015/EKU (0180/15)) and Informed Patient Consents at the Department of Pulmonology in the Semmelweis Hospital (Miskolc, Hungary). For this pilot study, the samples were pooled in order to minimize information loss of species below the detection threshold and improve

efficiency following the method described in [26]. Serum samples from 90 healthy individuals (control), 90 lung cancer patients, 90 COPD patients and 90 patients with comorbidity of COPD with lung cancer were separately pooled.

Preparation of human serum samples included denaturation, glycan release, fluorophore labeling and magnetic bead mediated cleanup. First, 2 µL of serum samples were diluted with HPLC grade water to 10 µL. Since the Fast Glycan Sample Preparation and Analysis protocol was optimized for purified IgG samples, to avoid possible precipitation issues with the more complex serum samples, a modified denaturation protocol was used by adding 5 µL denaturation solution and applying 40℃ for 10 min followed by 70℃ for 10 min. The glycan release process was also modified and performed by the addition of 1.0 µL of PNGase F enzyme (200 mU) to the reaction mixture and incubated at 60℃ for 1 hour instead of 20 min, to ensure complete removal of the serum *N*-glycome. The endoglycosidase digestion reaction was stopped by the addition of 120 µL of ice-cold acetonitrile. This was followed by drying the reaction mixture under reduced pressure at 60℃ for 1 hour in a SpeedVac. The dried samples were reconstituted in the labeling solution containing 4.0 µL of 40 mM 8-aminopyrene-1,3,6-trisulfonic acid (APTS) in 20% acetic acid, 2.0 µL of $NaBH_3CN$ (1 M in THF) and 4 µL 20% acetic acid. The reaction mixture was incubated in a heating block using a modified evaporative labeling protocol with closed vial cap at 50°C for 60 min, followed with open cap at 55°C for 80 minutes [27]. After the labeling step, the samples were purified by magnetic beads following the Fast Glycan Sample Preparation and Analysis protocol and analyzed by CE-LIF. Exoglycosidase digestions were performed by consecutive additions of sialidase A to remove all α 2-3,6,8-linked sialic acids, Jack bean galactosidase to remove β1-4,6-linked galactose residues and Jack bean hexosaminidase to remove the β1-2,4,6-

linked *N*-acetyl-glucosamines by respective overnight incubations at 37°C as described earlier in [28].

Capillary electrophoresis

Capillary electrophoresis analyses with laser induced fluorescent detection (CE-LIF) were performed using a PA800 Plus Pharmaceutical Analysis System (SCIEX). All CE measurements were accomplished in 40 cm effective length (50 cm total length), 50 µm ID bare fused silica capillaries filled with the HR-NCHO separation gel buffer (SCIEX). 30 kV electric potential was applied during the separation step in reversed polarity mode (cathode at the injection side, anode at the detection side) at 30°C. To increase detection sensitivity and reproducibility, a three-stage sample injection procedure was used: Step 1) 1.0 psi for 5.0 sec water pre-injection, Step 2) 3.0 kV for 3.0 sec sample injection and Step 3) 2.0 kV for 2.0 sec bracketing standard injection. This latter was used for high precision GU value determination by the GUcal software (www.GUcal.hu) [29]. Data collection and analysis were implemented by the 32Karat (version 10.1) software package (SCIEX). Relative percentage area values of the separated peaks were calculated by the PeakFit v4.12 Software (SeaSolve Software Inc., San Jose, CA).

Results and Discussion

In this pilot study, a new sample preparation method was applied to accommodate the high complexity of the human serum samples including temperature adjusted denaturation and extended glycan release and evaporative labeling protocol. Capillary electrophoresis - laser induced fluorescence detection was used to analyze and compare the *N*-glycosylation patterns of pooled human serum samples from 90 patients each with chronic obstructive pulmonary disease (COPD),

lung cancer (LC), and their comorbidity (COPD with LC) to healthy individuals (as control). For better comparability and consequently easier structural identification, the timescale of the acquired electropherograms were converted from migration time to GU values using the GUcal software (freely available from GUcal.hu) [29]. A representative electropherogram of the pooled healthy human serum sample *N*-glycome is shown in Figure 1, featuring the separation of 61 peaks. Structural elucidation of all separated *N*-glycans utilized direct mining of GU database entries (GlycoStore.org), exoglycosidase digestion based carbohydrate sequencing [28], comparison to oligosaccharides released from carefully chosen glycoprotein standards (ribonuclease B, immunoglobulin G, α-1-antitrypsin), and some earlier published literature data on the same subject matter [30-32]. The exoglycosidase based glycan sequencing process is shown in Figure 2, utilizing sialidase A, galactosidase and hexosaminidase, depicted by the corresponding traces. Sequence information was derived from the GU value shifts of the individual peaks as the result of the consecutive exoglycosidase treatments [33].

Table 1 lists all identified *N*-glycan structures in the pooled human serum sample as numbered in Figure 1. The first level of structural elucidation of the separated glycans accomplished by using their GU values to search the GlycoStore database (glycostore.org). Glycans denoted by # were identified considering the results of the sequential exoglycosidase digestion process shown in Figure 2. Structure identification of entries marked with asterisks were accomplished based on a comparative exercise utilizing the *N*-glycan profiles of IgG (*), RNase B (**) and AAT (***) [32, 34, 35]. It is important to note that besides these 61 carbohydrates, no other glycan structures were detected in the pathological samples compared to the control.

After the identification of the separated *N*-glycan structures in the pooled healthy human control serum sample, the peak areas of all peaks were computed and quantitatively compared to lung cancer, COPD and their comorbidity (COPD with LC) pooled sample data with their respective SD values based on the triplicate runs. Before the integration step for peak area determination, all electropherograms were normalized to Peak #61 (FA2BG2), i.e., the RFU values were divided by the RFU value of Peak #61, because it was apparently stable in size in all runs and well-separated from the other peaks. This normalization step was necessary to adequately calculate the peak areas of some of the not completely separated peaks. After the normalization step, the relative peak area % values were calculated by dividing the integrated peak areas by the sum of all integrated areas and multiplied by 100. Table 1 lists the suggested glycan structures corresponding to the separated peaks in Figure 1, the calculated CE-LIF GU values and relative peak areas with their SD. Only peaks with >1% relative peak areas were taken into consideration in this comparative pilot study (highlighted bold in Table 1). All runs were done in triplicates and the average RSD of the relative peak areas was 3.46%. Please note that rows 62 - 78 in Table 1 represent the peaks appeared during exoglycosidase digestion after each sequential analysis step.

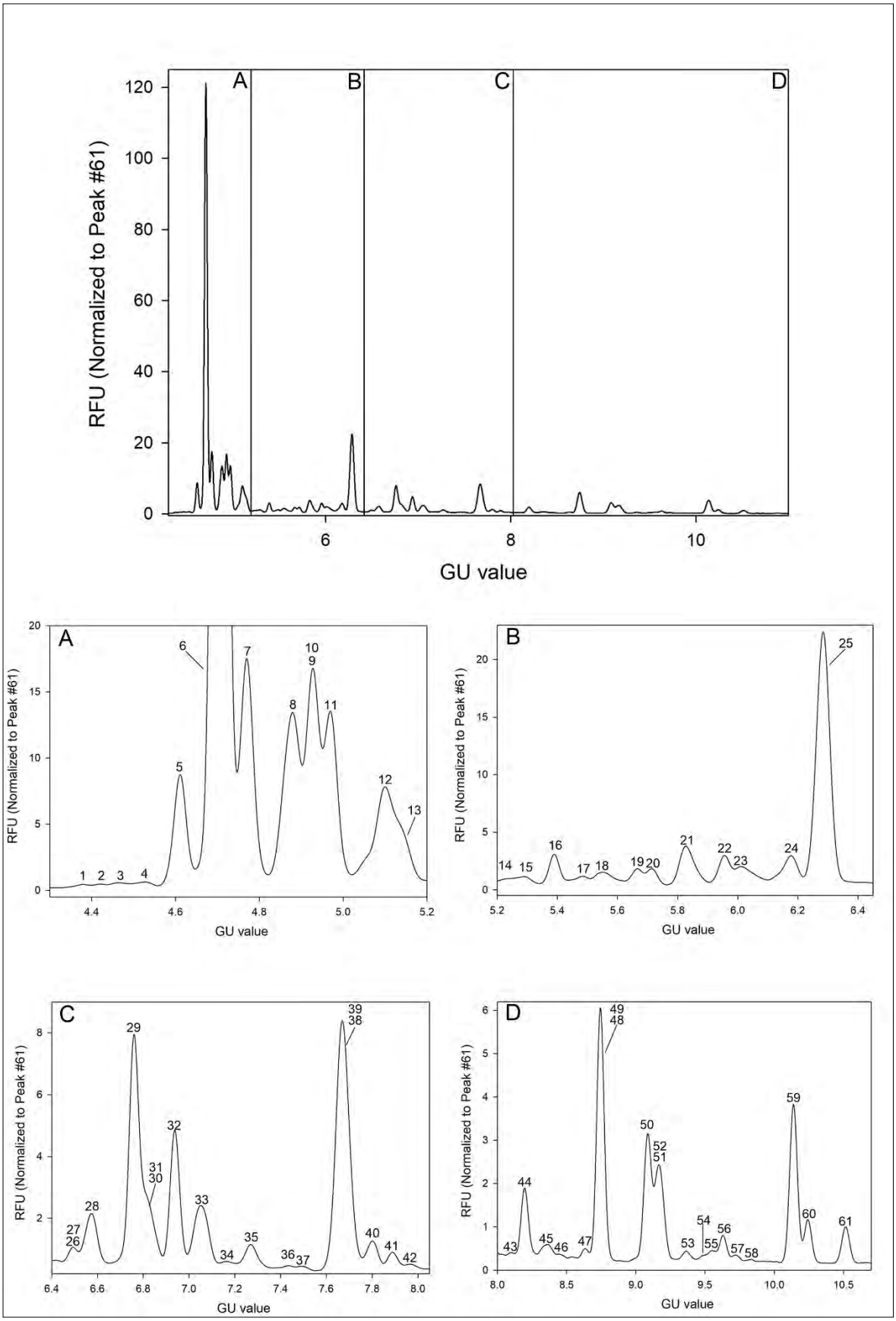

A
B
C
D
RFU (Normalized to Peak #61)
GU value
1 2 3 4
5
6
7
8
9 10
11
12
13
14 15
16
17 18
19 20
21
22
23
24
25
26 27
28
29
30 31
32
33
34
35
36 37
38 39
40
41
42
43
44
45 46
47
48 49
50
51 52
53
54 55
56
57 58
59
60
61

Figure 1. Capillary electrophoresis analysis of the released and APTS labeled *N*-glycans from pooled healthy human serum (upper panel). The lower panels show the enlarged sections of the electropherogram: A) GU 4.3-5.2; B) GU 5.2-6.4; C) GU 6.4-8.1; and D) GU 8.0-10.7. The corresponding structures, GU values and relative peak areas with their SD are listed in Table 1. Separation conditions: 40 cm effective capillary length (50 cm total length), 50 µm ID bare fused silica; 30 kV (0.17 min ramp time) separation voltage in reversed polarity mode. LIF detection (excitation: 488 nm/emission: 520 nm); separation temperature 30 °C. Injection: water pre-injection 5.0 s at 1.0 psi, followed by 3.0 kV/3.0 s sample and 2.0 kV/2.0 s bracketing standard (DP2 +DP15).

Table 1. The identified *N*-glycan structures in the pooled healthy control human serum sample, their CE-LIF GU values and relative peak areas with SD. Bold entries represent peaks with >1% relative peak areas (at least in one sample groups from control, COPD, LC and LC with COPD). Entries from 62 to 78 depict the glycan structures of the sequencing analysis results (shown in Figure 2). Abbreviated structure names followed the nomenclature suggested in [36]. Entries were identified as follows: 1) based on their GU values using the GlycoStore database (glycostore.org); 2) considering the results of the sequential exoglycosidase digestion experiments shown in Figure 2 (denoted with #) and 3) all other glycan structures were identified based on *N*-glycosylation information of high abundant glycoproteins in serum: IgG (*), RNase B (**) and AAT (***).

	N-glycan structure	GU	Relative peak area (%)			N-glycan structure	GU	Relative peak area (%)	
1	A4G4[3,3,3,6]S4#	4.4	0.0955	±0.003	40	M6**	7.8	0.8338	±0.039
2	A4G4[3,3,3,3]S4#	4.46	0.0935	±0.002	41	A3G3[3]S1#	7.89	0.5296	±0.013
3	A1G1S1#	4.51	0.1307	±0.006	42	FA3#	7.98	0.4379	±0.011
4	FA4G4[3,3,3,6]S4***	4.57	0.1285	±0.004	43	A2B[6]G1	8.08	0.2962	±0.009
5	**FA4BG4[3,3,3,3]S4#**	**4.62**	**0.6814**	**±0.015**	**44**	**FA2B***	**8.2**	**1.9016**	**±0.064**
6	**A2G2[6]S2***	**4.71**	**21.8425**	**±0.953**	**45**	A2B[3]G1*	8.35	0.6615	±0.026
7	**FA3G3[6]S3#**	**4.77**	**0.9032**	**±0.026**	46	M7**	8.48	0.3190	±0.014
8	**A2G2[3]S2*****	**4.88**	**2.4776**	**±0.098**	47	M7**	8.62	0.3836	±0.013
9	**A2BG2S2***	**4.93**	**1.5508**	**±0.038**	**48**	**FA2[6]G1***	**8.75**	**6.1207**	**±0.293**
10	**M3**				**49**	**M7****			
11	**FA2G2S2***	**4.97**	**2.3546**	**±0.052**	**50**	**FA2[3]G1***	**9.09**	**2.9771**	**±0.134**
12	**FA2BG2S2***	**5.1**	**1.2156**	**±0.026**	**51**	**FA2B[6]G1***	**9.17**	**2.5049**	**±0.067**
13	**FA3G3[3]S3*****	**5.14**	**1.2645**	**±0.031**	**52**	**M8****			
14	A2[6]G1S1*	5.24	0.1604	±0.007	53	A4G4[6]S1	9.34	0.1668	±0.008
15	A2[3]G1[6]S1	5.3	0.1631	±0.005	54	A4G4[3]S1	9.48	0.1636	±0.004
16	FA3BG3S3#	5.39	0.5907	±0.026	55	FA2B[3]G1*	9.55	0.4411	±0.012
17	**A4G4[6]S3**	**5.47**	**0.8155**	**±0.040**	56	M8**	9.64	0.5983	±0.023
18	A2[3]G1[3]S1*	5.56	0.2984	±0.012	57	M8**	9.73	0.3001	±0.009
19	FA2G2[6]S2***	5.66	0.3043	±0.018	58	FA4***	9.86	0.0874	±0.004

20	**A3G3[6]S2**	**5.72**	**1.1683**	**±0.031**	**59**	**FA2G2***	**10.14**	**3.5100**	**±0.089**
21	FA2[6] G1S1*	5.83	0.7706	±0.029	60	M9**	10.25	0.9882	±0.042
22	**A4G4[3,3,3]S3**	**5.95**	**0.9457**	**±0.024**	61	FA2BG2*	10.52	0.8847	±0.035
23	FA2[3]G1S1*	6.03	0.5180	±0.017	62	A1G1	6.62	Sequencing	
24	**A3G3[3]S2**	**6.17**	**0.8196**	**±0.030**	63	A2[3]G1	8.12	Sequencing	
25	**A2G2[6]S1***	**6.28**	**13.1137**	**±0.465**	64	A2G2	9.17	Sequencing	
26	FA2B[6]G1S1*	6.48	0.4828	±0.016	65	A2BG2	10	Sequencing	
27	FA2B[3]G1S1*				66	A3[6]G3	11.36	Sequencing	
28	**A2BG2S1***	**6.58**	**1.3189**	**±0.034**	67	A3[3]G3	11.63	Sequencing	
29	**FA2G2S1***	**6.76**	**4.1359**	**±0.095**	68	FA3G3	12.33	Sequencing	
30	**A2***	**6.82**	**2.7863**	**±0.119**	69	FA3BG3	12.54	Sequencing	
31	**M5****				70	A4G4	13.72	Sequencing	
32	**FA2BG2S1***	**6.94**	**2.2966**	**±0.063**	71	FA4G4	14.38	Sequencing	
33	**A4G4[6]S2**	**7.05**	**1.3916**	**±0.049**	72	FA4BG4	14.45	Sequencing	
34	A4G4[3]S2	7.17	0.5179	±0.023	73	A1	5.86	Sequencing	
35	**A2B***	**7.27**	**0.7528**	**±0.023**	74	A3	7.81	Sequencing	
36	A2[6]G1	7.4	0.1440	±0.004	75	A4	8.94	Sequencing	
37	A3G3[6]S1	7.51	0.1470	±0.005	76	FA3B	9.93	Sequencing	
38	**FA2***	**7.67**	**9.5132**	**±0.270**	77	FA4B	11.02	Sequencing	
39	**M6****				78	FM3	5.09	Sequencing	

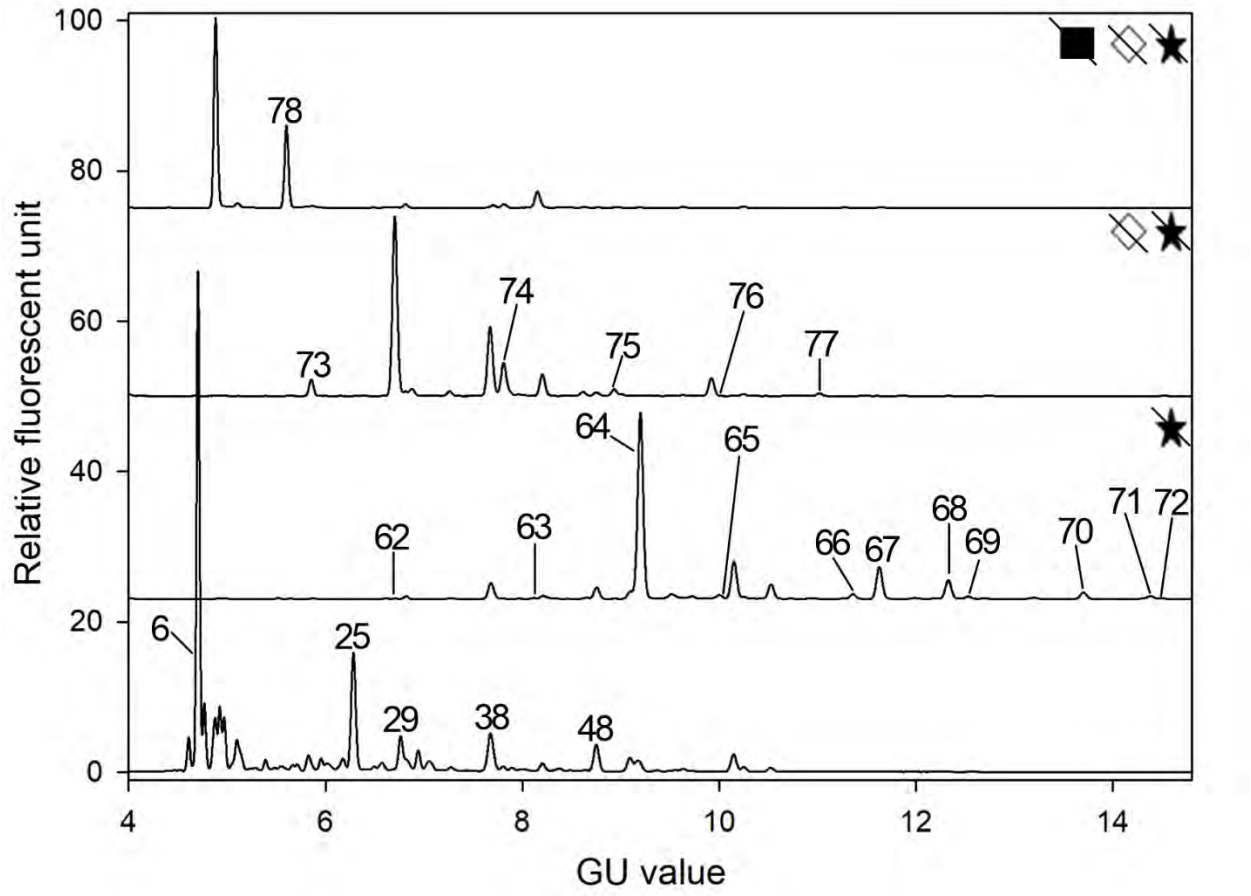


Figure 2. Exoglycosidase digestion based carbohydrate sequencing of the pooled healthy human control serum *N*-glycome using sialidase (★), galactosidase (◇), and hexosaminidase (■) to verify the identity of the glycan structures of interest. Separations conditions were the same as in Figure 1. The corresponding structures and GU values are listed in Table 1.

Changes in relative peak areas of the pooled lung cancer, COPD and comorbidity (COPD with LC) samples were compared to the pooled healthy control, based on their capillary electrophoresis analysis results. As shown in Table 2, only glycans with significant alterations were taken into consideration during the evaluation process, i.e., *N*-glycan structures satisfying the following two criteria: 1) the relative peak areas of the *N*-linked glycan structures were >1%, at least in one sample group of healthy control, COPD, lung cancer, or their comorbidity (bold structures in Table

1); and 2) at least one of the observed relative peak area differences between any of the disease groups and the control was >33%.

As one can observe, several relative peak area values in the pooled LC sample were considerably altered compared to the control (Table 2, Column 3), such as FA4BG4[3,3,3,3]S4 (peak 5: +167%±7.4), FA3G3[6]S3 (peak 7: +176%±9.7), A2BG2S2 + M3 (peaks 9-10: +133%±6.3), FA2G2S2 (peak 11: 105%±5.1), FA2BG2S2 (peak 12: 50%±2.9), A2BG2S1 (peak 28: -33%±1.6), FA2G2S1 (peak 29: -41%±2.1), FA2BG2S1 (peak 32: -41%±2.1), A2B (peak 35: +36%±2.1) and FA2G2 (peak 59: -36%±1.7). These changes may have diagnostic glycobiomarker potential for lung cancer. On the other hand, one should take into consideration that the following relative peak areas also increased in COPD: FA4BG4[3,3,3,3]S4 (peak 5, LC:+167%±7.4, COPD: +106%±5.2), A2BG2S2 + M3 (peaks 9-10, LC:+133%±6.3, COPD: +58%±2.8) and FA2BG2S2 (peak 12, LC:+50%±2.9, COPD: +33%±1.8), therefore, they may indicate the presence of both lung cancer or COPD, but not the comorbidity. In addition the increment of the relative peak area of FA3G3[6]S3 (peak 7, LC:+176%±9.7, COPD: +79%±3.8, LC withCOPD: +56%±3.0) could represent either lung cancer or COPD or their comorbidity. Interestingly, peak 28, corresponding to the A2BG2S1 structure slightly increased (+12%±0.6) in comparison of COPD to control, while significantly decreased (-33%±1.6) in lung cancer compared to control, therefore could also be a glycobiomarker for lung cancer. Moreover, using Column 4 in Table 2, one can distinguish COPD from lung cancer considering the increasing changes in FA3G3[6]S3, A2BG2S2+ M3, FA2G2S2, A2B and FA2 + M6 (54%±2.5, 47%±2.2, 79%±4.2, 35%±2.0 and 41%±2.1), and the decreasing values of A2BG2S1 and FA2G2S1 (41%±1.9 and 39%±2.0).

In addition, the relative peak area change of peak 29 (FA2G2S1) may discern comorbidity of COPD with lung cancer (-37 %±1.7) from COPD (-4%±0.2), but not from lung cancer (-41%±2.1). In all pathological samples peak 59 (FA2G2, LC: -36%±1.7, COPD: -34%±1.6, LC withCOPD: -39%±2.1) decreased, thus, possibly distinguish them only from the control. Please note that FA2 + M6 (peaks 38-39) and A2BG2S2 + M3 (peaks 9-10) structures were co-migrating, thus their individual alterations could not be precisely evaluated.

Table 2. Comparison of the relative peak areas of lung cancer (LC), COPD, their comorbidity (COPD with LC) to the control sample and between LC and COPD, where at least one observed difference was >33% (bold numbers) for any peak with >1% relative area between the disease and control samples.

Peak Number	N-Glycan structure	LC to Control %	COPD to Control %	LC + COPD to Control %	LC to COPD %
5	FA4BG4[3,3,3,3]S4	**167 ±7.4**	**106 ±5.2**	18 ±0.9	29 ±1.4
7	FA3G3[6]S3	**176 ±9.7**	**79 ±3.8**	**56 ±3.0**	**54 ±2.5**
9-10	A2BG2S2, M3	**133 ±6.3**	**58 ±2.8**	29 ±1.4	**47 ±2.2**
11	FA2G2S2	**105 ±5.1**	15 **±0.7**	4 ±0.2	**79 ±4.2**
12	FA2BG2S2	**50 ±2.9**	**33 ±1.8**	16 ±0.7	13 ±0.9
28	A2BG2S1	**-33 ±1.6**	12 ±0.6	-17 ±1.0	**-41 ±1.9**
29	FA2G2S1	**-41 ±2.1**	-4 ±0.2	**-37 ±1.7**	**-39 ±2.0**
32	FA2BG2S1	**-41 ±2.1**	-30 ±1.5	-26 ±1.3	-17 ±0.8
35	A2B	**36 ±2.1**	0 ±0	12 ±0.6	**35 ±2.0**
38-39	FA2, M6	15 ±0.9	-18 ±0.9	28 ±1.6	**41 ±2.1**
59	FA2G2	**-36 ±1.7**	**-34 ±1.6**	**-39 ±2.1**	-3 ±0.1

Figure 3 shows the quantitative alterations among the relative peak areas within the *N*-glycan subclasses specified by branching, degrees of sialylation and total fucosylation, determined in the control, COPD, lung cancer and comorbidity of COPD with lung cancer patient serum samples. Variations in the size of the carbohydrate structures were evaluated based on their degree of

branching, i.e., the amounts of various mono-, bi-, tri-, and tetra-antennary types. The number of terminal sialic acids was also taken into consideration in the evaluation process, focusing on the tendency of the relative peak area changes in mono-, bi-, tri- and tetra-sialylation. The relative peak areas of mono-sialylated glycans slightly increased in COPD (+5%±0.2), slightly decreased in comorbidity of COPD with lung cancer (-6%±0.2) and more substantially decreased in lung cancer (-27%±0.9). Moreover, the relative peak areas of bi-, tri- and tetrasialylated subclasses variably increased in COPD, lung cancer and their comorbidity (COPD with LC) in comparison to the pooled control healthy human serum and the increment was the highest in lung cancer (bi-antennary: +12%±0.4, tri-antennary: +42%±1.4, tetra-antennary: +168%±5.8), medium in COPD (bi-antennary: +7%±0.2, tri-antennary: +16%±0.6, tetra-antennary: +117%±4.1) and the smallest in their comorbidity (bi-antennary: +6%±0.2, tri-antennary: +8%±0.3, tetra-antennary: +23%±0.8). The relative peak areas of fucosylated *N*-glycans slightly decreased in COPD (-8%±0.3) and also in comorbidity of COPD with lung cancer (-4%±0.1) in contrast to lung cancer, where a small increase (+7%±0.2) was observed.

The relative peak areas of mono-, tri- and tetra-antennary glycans similarly to the bi-, tri- and tetra-sialylated subclasses increased in all three disease groups compared to the healthy control with the highest increment in lung cancer (mono-antennary: +116%±4.0, tri-antennary: +45%±1.6, tetra-antennary: 44.5%±1.5), medium in COPD (mono-antennary: +59%±2.0, tri-antennary: +15%±0.5, tetra-antennary: 31%±1.1) and the smallest in their comorbidity (mono-antennary: +30%±1.0, tri-antennary: +7.5%±0.3, tetra-antennary: 3%±0.1). In contrary, the relative peak areas of bi-antennary *N*-glycans hardly decreased or not changed at all in all sample types of COPD (-3%±0.1), lung cancer (-5%±0.2) and comorbidity of COPD with lung cancer (-0.3%±0.0),

compared to the control. However, taking a closer look at the data, these observed changes were apparently caused by the interplay between various increasing and decreasing amounts of bi-antennary structures, i.e., the relative peak areas of some bi-antennary glycans significantly decreased (e.g., A2BG2S1 (peak 28): LC: -33%±1.6, LC withCOPD: -17%±1.0; FA2G2S1 (peak 29): LC: -41%±2.1, COPD: -4%±0.2, LC withCOPD: -37%±1.7; FA2G2 (peak 59): LC: -36%±1.7, COPD: -34%±1.6, LC withCOPD: -39%±2.1); while others increased (e.g., A2BG2S2 + M3) (peaks 9-10): LC: +133%±6.3, COPD: +58%±2.8, LC with COPD: +29%±1.4; FA2G2S2 (peak 11): LC: +105%±5.1, COPD: +15%±0.7, LC withCOPD: +4%±0.2; FA2BG2S2 (peak 12): LC: +50%±2.9, COPD: +33%±1.8, LC withCOPD: +16%±0.7 and A2B (peak 35): LC: +36%±2.1, LC withCOPD: +12%±0.6), thus their changes were almost counterbalanced. Based on these observations, as a first approximation we consider that *N*-glycan subclasses of mono- tri- and tetrasialylated, fucosylated as well as mono-, tri- and tetra-antennary forms could serve as potential biomarkers for lung cancer, COPD and their comorbidity.

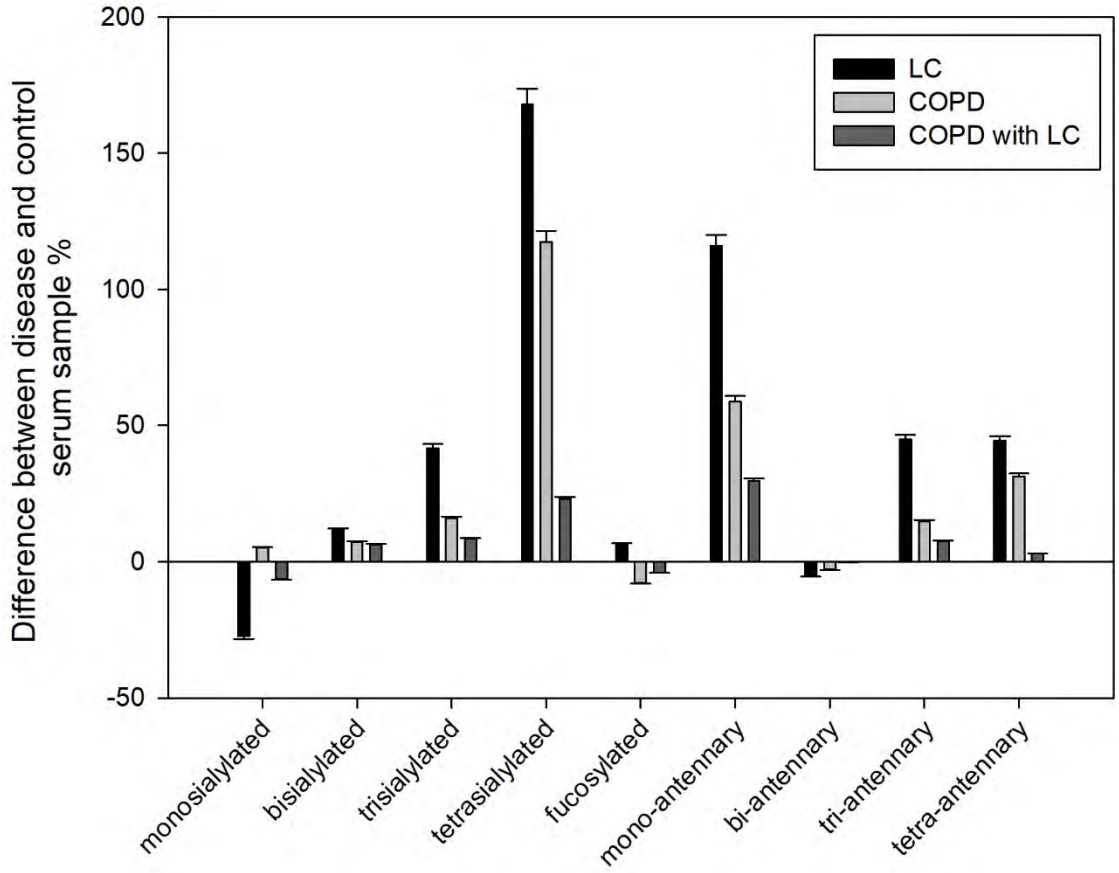


Figure 3. Alterations among the relative peak areas of specific *N*-glycan subclasses (Sialoforms: mono-, bi-, tri- and tetra-sialo; Fucosylated and Branching: mono-, bi- tri- and tetra-antennary) of lung cancer (black), COPD (gray) and their comorbidity (dark gray) with their corresponding RSDs. The results were calculated based on the data in supplementary Table 1.

Conclusion

In this pilot study, the *N*-glycosylation profiles of patient samples of chronic inflammatory (COPD) and malignant (LC) pulmonary diseases as well as their comorbidity (COPD with LC) were quantitatively studied and compared to healthy controls. A novel temperature adjusted denaturation protocol and extended enzymatic release and evaporative derivatization time was

used for the asparagine linked oligosaccharides from the complex serum samples, which were then analyzed by capillary electrophoresis with high sensitivity laser induced fluorescence detection. Sixty-one *N*-glycan structures were identified in the control human serum samples and since no other glycans appeared in any of the three disease categories, only these 61 structures were quantitatively monitored in this study. Our results suggested that certain serum *N*-glycans could be used as potential markers for the different types of pulmonary diseases. Therefore, the panel of the 13 glycans listed in Table 2 could be considered to differentiate lung cancer, COPD and their comorbidity from the control and LC from COPD. In addition, alterations in the *N*-glycan subclasses, such as fucosylated, mono-, bi-, tri- and tetra-sialylo, as well as mono-, bi-, tri- and tetra-antennary glycans could also carry interesting diagnostic information. The glycan panel in Table 2 and the corresponding subclasses may provide even more reliable information as they represent the sum of multiple structural changes caused by a given disease. This is especially applicable for the highly branched sialylated structures as our recent genotyping data suggested significant increase of MGAT5 activity, i.e., increased branching in lung cancer [37]. Currently we are in the process of collecting 300 samples from each the disease groups that will be individually analyzed in view of our preliminary pooled sample based results.

Acknowledgement

The authors gratefully acknowledge the support of the National Research, Development and Innovation Office (NKFIH) (K 116263) grants of the Hungarian Government. This work was also supported by the BIONANO_GINOP-2.3.2-15-2016-00017 project, the V4-Korea Joint Research Program, project National Research, Development and Innovation Office (NKFIH) (NN 127062) grants of the Hungarian Government and the Human Capacities Grant Management Office, the ÚNKP-18-3-I-DE-393 New National Excellence Program Hungarian Ministry of Human

Capacities and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. This is contribution #152 from the Horvath Csaba Memorial Laboratory of Bioseparation Sciences.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin., 68 (2018) 394-424.

[2] A.L. Durham, I.M. Adcock, The relationship between COPD and lung cancer, Lung Cancer, 90 (2015) 121-127.

[3] J.L. López-Campos, W. Tan, J.B. Soriano, Global burden of COPD, Respirology, 21 (2016) 14-23.

[4] D. Spyratos, E. Papadaki, S. Lampaki, T. Kontakiotis, Chronic obstructive pulmonary disease in patients with lung cancer: prevalence, impact and management challenges, Lung Cancer (Auckl), 8 (2017) 101-107.

[5] Y. Wang, F. Jiang, X. Tan, P. Tian, CT-guided percutaneous transthoracic needle biopsy for paramediastinal and nonparamediastinal lung lesions: Diagnostic yield and complications in 1484 patients, Medicine, 95 (2016) e4460-e4460.

[6] K. Shyamala, H.C. Girish, S. Murgod, Risk of tumor cell seeding through biopsy and aspiration cytology, J. Int. Soc. Prev. Community Dent., 4 (2014) 5-11.

[7] Y. Xi, J. Fan, D. Che, K. Zhai, T. Ren, X. Feng, L. Shang, J. Hu, Y. Yu, Q. Meng, Distant Metastasis and Survival Outcomes after Computed Tomography-Guided Needle Biopsy in Resected Stage I-III Non-Small Cell Lung Cancer, J. Cancer, 8 (2017) 3356-3361.

[8] O. Gornik, T. Pavic, G. Lauc, Alternative glycosylation modulates function of IgG and other proteins - implications on evolution and disease, Biochim. Biophys.Acta, 1820 (2012) 1318-1326.

[9] T. Suzuki, K. Ohtsubo, N. Taniguchi, Sugar Chains: Decoding the Functions of Glycans, 2015.

[10] H. Lemjabbar-Alaoui, A. McKinney, Y.-W. Yang, V.M. Tran, J.J. Phillips, Chapter Nine - Glycosylation Alterations in Lung and Brain Cancer, in: R.R. Drake, L.E. Ball (Eds.) Advances in Cancer Research, Academic Press2015, pp. 305-344.

[11] F. Clerc, K.R. Reiding, B.C. Jansen, G.S. Kammeijer, A. Bondt, M. Wuhrer, Human plasma protein N-glycosylation, Glycoconj. J., 33 (2016) 309-343.

[12] A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, H.P. Packer, J.H. Prestegard, R.L. Schnaar, P.H. Seeberger, Essentials of Glycobiology, Cold Spring Harbor Laboratory Press, DOI (2015-2017).

[13] R.R. Townsend, A.T. Hotchkiss, Techniques in Glycobiology, Marcel Dekker Inc., New York, NY, 1997.

[14] R. Saldova, L. Royle, C.M. Radcliffe, U.M. Abd Hamid, R. Evans, J.N. Arnold, R.E. Banks, R. Hutson, D.J. Harvey, R. Antrobus, S.M. Petrescu, R.A. Dwek, P.M. Rudd, Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG, Glycobiology, 17 (2007) 1344-1356.

[15] T. Pavic, D. Dilber, D. Kifer, N. Selak, T. Keser, D. Ljubicic, A. Vukic Dugac, G. Lauc, L. Rumora, O. Gornik, N-glycosylation patterns of plasma proteins and immunoglobulin G in chronic obstructive pulmonary disease, J. Transl. Med., 16 (2018) 323.

[16] A. Bondt, S. Nicolardi, B.C. Jansen, K. Stavenhagen, D. Blank, G.S. Kammeijer, R.P. Kozak, D.L. Fernandes, P.J. Hensbergen, J.M. Hazes, Y.E. van der Burgt, R.J. Dolhain, M. Wuhrer, Longitudinal monitoring of immunoglobulin A glycosylation during pregnancy by simultaneous MALDI-FTICR-MS analysis of N- and O-glycopeptides, Sci. Rep., 6 (2016) 27955.

[17] A. Bondt, S. Nicolardi, B.C. Jansen, T.M. Kuijper, J.M.W. Hazes, Y.E.M. van der Burgt, M. Wuhrer, R.J.E.M. Dolhain, IgA N- and O-glycosylation profiling reveals no association with the pregnancy-related improvement in rheumatoid arthritis, Arthritis research & therapy, 19 (2017) 160-160.

[18] Y. Liang, T. Ma, A. Thakur, H. Yu, L. Gao, P. Shi, X. Li, H. Ren, L. Jia, S. Zhang, Z. Li, M. Chen, Differentially expressed glycosylated patterns of alpha-1-antitrypsin as serum biomarkers for the diagnosis of lung cancer, Glycobiology, 25 (2015) 331-340.

[19] T. Fujimura, Y. Shinohara, B. Tissot, P.C. Pang, M. Kurogochi, S. Saito, Y. Arai, M. Sadilek, K. Murayama, A. Dell, S. Nishimura, S.I. Hakomori, Glycosylation status of haptoglobin in sera of patients with prostate cancer vs. benign prostate disease or normal subjects, Int. J. Cancer, 122 (2008) 39-49.

[20] C. Varadi, S. Mittermayr, A. Szekrenyes, J. Kadas, L. Takacs, I. Kurucz, A. Guttman, Analysis of haptoglobin N-glycome alterations in inflammatory and malignant lung diseases by capillary electrophoresis, Electrophoresis, 34 (2013) 2287-2294.

[21] O. Choi, N. Tomiya, J.H. Kim, J.M. Slavicek, M.J. Betenbaugh, Y.C. Lee, N-glycan structures of human transferrin produced by Lymantria dispar (gypsy moth) cells using the LdMNPV expression system, Glycobiology, 13 (2003) 539-548.

[22] A. Quaranta, A. Sroka-Bartnicka, E. Tengstrand, G. Thorsén, N-Glycan profile analysis of transferrin using a microfluidic compact disc and MALDI-MS, Analytical and bioanalytical chemistry, 408 (2016) 4765-4776.

[23] E. Ito, R. Oka, T. Ishii, H. Korekane, A. Kurimoto, Y. Kizuka, S. Kitazume, S. Ariki, M. Takahashi, Y. Kuroki, K. Kida, N. Taniguchi, Fucosylated surfactant protein-D is a biomarker candidate for the development of chronic obstructive pulmonary disease, J. Proteomics, 127 (2015) 386-394.

[24] J.N. Arnold, R. Saldova, M.C. Galligan, T.B. Murphy, Y. Mimura-Kimura, J.E. Telford, A.K. Godwin, P.M. Rudd, Novel Glycan Biomarkers for the Detection of Lung Cancer, J. Proteome Res., 10 (2011) 1755-1764.

[25] L.R. Ruhaak, U.T. Nguyen, C. Stroble, S.L. Taylor, A. Taguchi, S.M. Hanash, C.B. Lebrilla, K. Kim, S. Miyamoto, Enrichment strategies in glycomics-based lung cancer biomarker development, Proteomics Clin. Appl., 7 (2013) 664-676.

[26] E.F. Schisterman, A. Vexler, To pool or not to pool, from whether to when: applications of pooling to biospecimens subject to a limit of detection, Paediatr Perinat Epidemiol, 22 (2008) 486-496.

[27] B. Reider, M. Szigeti, A. Guttman, Evaporative fluorophore labeling of carbohydrates via reductive amination, Talanta, 185 (2018) 365-369.

[28] A. Guttman, K.W. Ulfelder, Exoglycosidase matrix-mediated sequencing of a complex glycan pool by capillary electrophoresis, J. Chromatogr. A, 781 (1997) 547-554.

[29] G. Jarvas, M. Szigeti, A. Guttman, GUcal: An integrated application for capillary electrophoresis based glycan analysis, Electrophoresis, 36 (2015) 3094-3096.

[30] R. Saldova, A. Asadi Shehni, V.D. Haakensen, I. Steinfeld, M. Hilliard, I. Kifer, A. Helland, Z. Yakhini, A.L. Borresen-Dale, P.M. Rudd, Association of N-glycosylation with breast carcinoma and systemic features using high-resolution quantitative UPLC, J. Proteome Res., 13 (2014) 2314-2327.

[31] L. Royle, M.P. Campbell, C.M. Radcliffe, D.M. White, D.J. Harvey, J.L. Abrahams, Y.-G. Kim, G.W. Henry, N.A. Shadick, M.E. Weinblatt, D.M. Lee, P.M. Rudd, R.A. Dwek, HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software, Analytical Biochemistry, 376 (2008) 1-12.

[32] K.J. Lee, S.M. Lee, J.Y. Gil, O. Kwon, J.Y. Kim, S.J. Park, H.S. Chung, D.B. Oh, N-glycan analysis of human alpha1-antitrypsin produced in Chinese hamster ovary cells, Glycoconj. J., 30 (2013) 537-547.

[33] M. Szigeti, A. Guttman, Automated N-Glycosylation Sequencing Of Biopharmaceuticals By Capillary Electrophoresis, Sci. Rep., 7 (2017) 11663.

[34] A. Guttman, S. Herrick, Effect of the quantity and linkage position of mannose (alpha 1,2) residues in capillary gel electrophoresis of high-mannose-type oligosaccharides, Anal. Biochem., 235 (1996) 236-239.

[35] S. Mittermayr, J. Bones, A. Guttman, Unraveling the glyco-puzzle: glycan structure identification by capillary electrophoresis, Anal. Chem., 85 (2013) 4228-4238.

[36] D.J. Harvey, A.H. Merry, L. Royle, M.P. Campbell, P.M. Rudd, Symbol nomenclature for representing glycan structures: Extension to cover different carbohydrate types, Proteomics, 11 (2011) 4291-4295.

[37] Z. Elek, Z. Kovács, G. Keszler, M. Szabo, E. Csanky, J. Luo, A. Guttman, Z. Rónai, High throughput multiplex SNP-analysis in chronic obstructive pulmonary disease and lung cancer, Current Molecular Mecicine, 19 (2019) In Press.

ELSEVIER

Conflict of Interest Policy

Manuscript number (if applicable): Article Title: Comparative analysis of the human serum N-glycome in lung cancer, COPD and their comorbidity using capillary electrophoresis.	Corresponding Author name: Brigitta Mészáros

Declarations

Journal of Chromatography B requires that all authors sign a declaration of conflicting interests. If you have nothing to declare in any of these categories then this should be stated.

Conflict of Interest

A conflicting interest exists when professional judgement concerning a primary interest (such as patient's welfare or the validity of research) may be influenced by a secondary interest (such as financial gain or personal rivalry). It may arise for the authors when they have financial interest that may influence their interpretation of their results or those of others. Examples of potential conflicts of interest include employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/registrations, and grants or other funding.

Please state any competing interests

The authors do not have any conflict of interest.

Funding Source

All sources of funding should also be acknowledged and you should declare any involvement of study sponsors in the study design; collection, analysis and interpretation of data; the writing of the manuscript; the decision to submit the manuscript for publication. If the study sponsors had no such involvement, this should be stated.

Please state any sources of funding for your research

The authors gratefully acknowledge the support of the National Research, Development and Innovation Office (NKFIH) (K 116263) grants of the Hungarian Government. This work was also supported by the BIONANO_GINOP-2.3.2-15-2016-00017 project, the V4-Korea Joint Research Program, project National Research, Development and Innovation Office (NKFIH) (NN 127062) grants of the Hungarian Government and the Human Capacities Grant Management Office, the ÚNKP-18-3-I-DE-393 New National Excellence Program Hungarian Ministry of Human Capacities

Signature (a scanned signature is acceptable, but each author must sign)	**Print name**
	___András Guttman_____
Mészáros Brigitta	____Brigitta Mészáros _____
	____Márton Szigeti _____
	____Gábor Járvás _____

Kovács Zsuzsanna ____Zsuzsanna Kovács _____

Farkas ____Anna Farkas _____

Csánky Eszter ____Eszter Csánky _____

____Miklós Szabó _____

Renáta Kun ____Renáta Kun _____