SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Research of new lung cancer glycobiomarkers with a comprehensive approach

by Brigitta Bakos-Mészáros

Supervisor: András Guttman, Ph.D.

UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF MOLECULAR MEDICINE

DEBRECEN, 2023

Research of new lung cancer glycobiomarkers with a comprehensive approach

By Brigitta Bakos-Mészáros, Chemical Engineer MSc

Supervisor: András Guttman, PhD, DSc

Doctoral School of Molecular Medicine, University of Debrecen

Head of the **Defense Committee**:	László Csernoch, PhD, DSc
Reviewers:	Ferenc Csaba Budán, PhD
	János Kádas, PhD
Members of the Defense Committee:	Dezső Boda, PhD, DSc
	János Kerékgyártó, PhD

The PhD Defense takes place at the Lecture Hall of Building A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen, 29th of January, 2024, 13:00 p.m.

1. Introduction

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 reliable methods capable to readily diagnose them. Moreover, COPD increases the risk of lung cancer development, leading to their comorbidity. Therefore, there is an urgent need to develop novel molecular diagnostic tools capable of predicting the presence and prognosis of the actual disease (LC, COPD or comorbidity) with adequate specificity and sensitivity. Therefore, new possibilities for non-invasive or partially invasive diagnostic procedures are widely investigated. The identification of molecular biomarkers in serum samples could be the basis for a diagnostic method based on blood testing. Biomarkers are specific objective indicators for biological or pathogen processes or therapeutic treatment [1]. The surface of living cells is covered by a large number of diverse complex sugars, which could be specifically modified even in the early stages of diseases. Malignant and inflammatory processes, such as lung cancer and COPD, could induce changes in glycan biosynthesis, so they could result in varioius differences in the post-translational modification of the cell surface and circulating glycoproteins in blood. The identification of these specific modifications could be used for a diagnostic purposes. In the last decades, research and application of glycobiomarkers are getting more and more attention in the development of diagnostic procedures [2]. The reliable and validable identification of molecular alterations associated with certain diseases often requires the processing of a large number of samples, which also increases the necessity for the development of new statistical methods.

2. Objective

Late recognition is one of the main reasons for the high death rate of lung cancer and one of the the simplest solutions could be early diagnosis. The aim of my PhD work was to endorse the diagnosis of lung cancer at an early stage through identification of potential glycobiomarkers. First, I intended to determine the asparagine-linked glycan profile of human blood serum by capillary electrophoresis equipped with a laser-induced fluorescence detector. After that, I aimed to identify the separated *N*-linked oligosaccharide peaks on the resulting electropherogram. The next goal was the comparison of the glycan profiles of healthy, lung cancer, COPD, as well as lung cancer and COPD patients in order to reveal significant specific differences. The identification of potential patient-specific glycobiomarkers could serve as a solid base for new molecular diagnostic method developments.

My second goal was the analysis of the samples of patients whose treatments were followed up. One sample set included lung cancer patients who had undergone resection surgery. The samples were taken before and after operation. The second sample set included lung cancer patients who had received chemotherapy. Samples were taken in different stages of their treatment. In both cases, my aim was to examine the differences in the human serum glycan profile before and after resection surgery and during chemotherapy, whether there was a correlation between clinical parameters and changes in glycan levels. Not only individual changes of certain glycans, but the combined alteration of several structures were investigated. Considering the results, the analysis of blood serum samples might give a potential prognostic picture of the outcome of the lung cancer patients' treatment, which could further increase the effectiveness of the therapy.

3. Results and discussion

Identification of *N*-linked glycans in human blood serum

During my thesis work I identified 61 asparagin linked glycan structures in healthy human blood serum. A novel temperature adjusted denaturation protocol as well as extended enzymatic release and evaporative derivatization time was used for preparing asparagine linked oligosaccharides from the complex serum samples. Measurements were performed by capillary electrophoresis equipped with high-sensitivity laser-induced fluorescence detection (CE-LIF). The identified oligosaccharide structures may contribute to the development of new glycobiomarker-based diagnostic methods.

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

I established a glycan panel that contained carbohydrate structures that could be potential glycobiomarkers for COPD, lung cancer and their comorbidity. In addition, alterations in *N*-glycan subclasses, such as fucosylated, mono-, bi-, tri- and tetrasialylated, as well as mono-, bi-, tri- and tetra-antennary glycans could also carry potential diagnostic information. The glycan panel and the corresponding subclasses may provide even more reliable information as they represent the sum of multiple structural changes caused by a given disease.

Investigation of the effect of resection surgery on the *N*-linked glycan profile of human serum

In my research, I investigated the changes in the *N*-linked glycans of serum proteins before and after lung tumor resection surgery. The samples were analyzed by CE-LIF. Despite the relatively small number of samples, I found significant correlations between changes in the amount of glycan structures and discrete and continuous clinical parameters. Analysis of the classification tree resulted in a set of *N*-glycan decision models that could even be used to monitor resection surgery of lung tumor. The positive outcome of the surgery showed significant correlation with the *N*-glycome change in patients who were still smoking and those who had already quit smoking, as well as in patients with and without atherosclerosis. In addition, surgeries with a negative outcome can be associated with a significant change in the relative peak area of certain *N*-glycan structures (FA2BG2S1, A2G2S(3)2, FA2[3]G1) both in the case of smokers and patients who have already quit smoking. Resection surgery with a negative outcome when lung cancer comorbidity with COPD or when atherosclerosis was not observed with lung cancer can be associated with changes in the relative peak area of

FA3G3S(6)3 and FA2BG2S1 or FA2BG2S2 and FA3G3S(3)3. Lung cancer comorbid with COPD and with or without atherosclerosis also showed significant correlation, with changes in the relative peak area of FA4BG4S(3)4 and FA2G2 or FA2G2S1. In patients, who had lung cancer comorbidity with COPD, a significant increase was observed due to its positive outcome in the relative peak area of FA2G2S2. In the *N*-glycan profile of human serum from lung cancer patients without atherosclerosis, who quit smoking resection surgery increased the relative peak area of the M9 glycan structure. The correlations listed above may provide a basis for a method to monitor the effectiveness of treatment during lung cancer resection surgery. After testing and validation, extension to a larger number of samples will provide an even more reliable and accurate picture of treatment efficacy.

In addition to analyzing the relationship between the clinical parameters and the changes in the relative peak area of each *N*-linked glycan structure, I also investigated the correlations between the changes in the relative peak area of groups of *N*-glycans and the clinical parameters. I determined the most significant clinical parameters that caused the change in the relative peak area of the entire afucosylated subgroup after the resection surgery. Lung cancer patients diagnosed with atherosclerosis, who quit smoking or had COPD, too, changed the amount of both total afucosylated and sialylated glycan structures in the same way. Patients who quit smoking, and had a positive outcome of the surgery, or did not have diabetes but had atherosclerosis, showed the same correlation in the change of the sialylated relative peak area. In case of lung cancer comorbidity with other diseases the correlation model showed not only the change of the relative peak area of the sialylated glycan group, but also the amount of the terminal galactosylated glycan group. In diabetic lung cancer patients, resection surgery generated a change in the relative peak area of the terminal galactosylated glycan group. Only one examined discrete clinical parameter (smoking patient with a positive outcome resection surgery) caused change in the relative peak area of neutral glycans with a satisfactory accuracy model.

I also evaluated the effect of continuous clinical parameters on the *N*-glycan profile by linear regression analysis. I found a linear correlation between the *N*-glycosylaiton changes in human serum caused by lung tumor resection surgery and continuous clinical parameters including age, CRP value, blood sugar level, the amount of cigarettes smoked, the years spent smoking and the stage of the disease. Using the correlations identified in this work, a panel of *N*-glycans could be assembled, from which the patient's medical history and condition can be deduced.

Examination of the effect of chemotherapy on changes in glycan pofile

I analyzed the correlations between the clinical parameters of lung cancer patients undergoing chemotherapy and the change of the asparagine-linked glycans due to chemotherapy. As a result of the analysis, changes in *N*-linked glycan structure due to chemotherapy can be modeled with sufficient reliability when the patient was not diagnosed with COPD just only with lung cancer (The glycan structures changed during chemotherapy: A3G3S(3)2, FA2B[6]G1, M8, M9, FA2[6]G1S1, FA2BG2S2, FA3G3S(3)3, FA2[3]G1, A2G2S(6)1, FA4BG4S(3)4 ΔA2G2S(6)2, ΔFA2G2S2 ΔFA2[6]G1. FA2, M6) In case of women, I also determined well-modeled changes depending on the treatment outcomes, tumor progression changed the amount of A3G3S(3)2 glycan structure and stagnant tumor state caused changes in the FA2G2S1, FA2, M6, FA2B, M9, FA2[6]G1 and M7 glycan structures. Diabetes comorbidity with COPD (FA3G3S(6)3, A2G2S(3)2) or atherosclerosis (M9, A2G2S(3)2) or without any of them (FA2[3]G1, FA2[6]G1, M7, FA2G2S1, M9) induced unique changes in the amount of the examined glycan structures after the chemotherapy. Lung cancer patients who did not have COPD either quit smoking (A3G3S(3)2, FA4BG4S(3)4, A2G2S(6)2, M9) or not (M9, FA2BG2S2, FA3G3S(3) 3, FA2[3]G1, A2G2S(6)1), the glycan structures determined in their human serum samples changed that can be followed by a specific model. The results also confirmed that changes in the sample of lung cancer diabetic patients the same model can be used in cases with tumor regression and without tumor progression (FA2G2S1, FA2G2S2, FA2[3]G1, M9). Reliable models were determined in patients who had atherosclerosis and lung cancer, when the tumor progressed after the treatment (A3G3S(3)2, A2G2S(6)1, FA2G2S2, A4G4S(6)2) or the tumor state did not change (A4G4S(6)2, A3G3S(3)2, FA2B[6]G1, M8, A2BG2S1, FA2BG2S2, FA3G3S(3)3, A2G2S(6)2). The reliability and accuracy of the obtained results could be increased by the investigation of more samples. The determined results in this work could help in the development of practical diagnostic methods monitoring the effectiveness of chemotherapy treatments.

In addition, the examined glycan structures were grouped into subclasses to analyze how the changes in the glycan subclasses after chemotherapy were related to the clinical parameters of the patients. The analysis resulted in well-defined models in female chemotherapy treated patients considering all three outcomes of chemotherapy (regression, progression, stagnant tumor). I determined reliable models for treatment-induced changes in patients with diabetes taking into account several clinical parameters, such as the without COPD, the with or without atherosclerosis, smoking or quitting smoking, tumor progression or tumor non-regression. Monitoring the changes in glycan subclasses increased the reliability of the results.

Unfortunately, due to the small number of samples, the classification tree analysis could not include the type of treatment, therefore the results were not sufficiently reliable. Linear regression analysis revealed insufficient correlation between chemotherapy treatments and continuous clinical parameters.

4. Summary

The main goal of my thesis work was to promote the research of *N*-glycan structures in blood serum and their applicability in diagnostic developments. Numerous studies have already proven that specific glycosylation processes take place as a result of malignant transformation and inflammations. I analyzed the *N*-glycan profile of human blood serum samples from lung cancer patients using CE-LIF and determined 61 *N*-glycan structures in human blood serum. These structures provided a glycan panel, in which I searched for potential glycobiomarkers for the diagnosis of lung cancer, COPD and their comorbidity. My goal was to find specific changes that were characteristic of the given disease or disease group. As a result, a panel of 13 *N*-linked glycan structures were determined containing potential glycobiomarkers for lung cancer, COPD and their comorbidity. The most significant changes were observed in the case of lung cancer affecting the following oligosaccharides (FA4BG4[3,3,3,3]S4, FA3G3[6]S3, A2BG2S2, M3, FA2G2S2, FA2BG2S2, A2BG2S1, FA2G2S1, FA2BG2S1, A2B, FA2G2). However, the amount of FA4BG4[3,3,3,3]S4, FA3G3[6]S3, A2BG2S2, M3, FA2BG2S2 and FA2G2 glycan structures also significantly changed in case of COPD. In case of lung cancer comorbidity with COPD the relative peak area of FA3G3[6]S3, FA2G2S1 and FA2G2 structures considerably changed. In addition to the individual examination of the specific glycan structures, I also investigated how lung cancer, COPD and their comorbidity affected the glycan subclasses. The development of lung cancer significantly altered the level of mono-tri- and tetrasialylated as well as the mono-tri- and tetra-antennary *N*-glycan structures. In case of COPD, smaller changes could be observed compared to lung cancer, however, the tetrasialylated, mono- and tetra-antennary classes had a great extent increment.

My second main goal was to reveal any correlation between changes in the *N*-glycan profile as a result of treatment and clinical parameters by examining the human serum samples of already diagnosed and treated patients. I analyzed the blood serum samples of lung cancer patients utilizing CE-LIF and applied a classification tree analysis in a novel way to process the data. My results proved that the method was suitable for mapping correlations between the changes of N-glycans and the clinical parameters in lung cancer patients. I analyzed two types of treatments, resection surgery and chemotherapy. In both, I found specific changes in the amount of the examined *N*-linked glycan structures as a result of the treatment in case of several clinical parameters. The specified correlations, after validation, could promote the development of a monitoring method measuring the effectiveness of surgeries and chemotherapy treatments and possibly improving the outcome of the treatments through timely interventions. Applying the

proper treatment plays an important role in effectiveness, and for selecting the right one, accurate diagnostic information is key.

[1] K. Strimbu, J.A. Tavel, What are biomarkers?, Curr. Opin. HIV AIDS. 5 (2010) 463–466. https://doi.org/10.1097/COH.0b013e32833ed177.

[2] S.A. Svarovsky, Cancer Glycan Biomarkers and their Detection - Past, Present and Future, Anal. Methods. (2006) 323–350.

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/24/2023.PL
Subject: PhD Publication List

Candidate: Brigitta Mészáros
Doctoral School: Doctoral School of Molecular Medicine

List of publications related to the dissertation

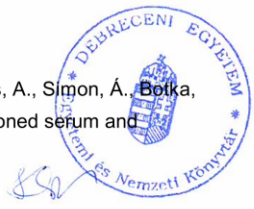
1. **Mészáros, B.**, Járvás, G., Farkas, A., Szigeti, M., Kovács, Z., Kun, R., Szabó, M., Csánky, E., Guttman, A.: Comparative analysis of the human serum N-glycome in lung cancer, COPD and their comorbidity using capillary electrophoresis.
   J. Chromatogr. B. 1137, 1-7, 2020.
   DOI: http://dx.doi.org/10.1016/j.jchromb.2019.121913
   IF: 3.205

2. **Mészáros, B.**, Járvás, G., Kun, R., Szabó, M., Csánky, E., Abonyi, J., Guttman, A.: Machine Learning Based Analysis of Human Serum N-glycome Alterations to Follow up Lung Tumor Surgery.
   Cancers (Basel). 12 (12), 1-13, 2020.
   DOI: http://dx.doi.org/10.3390/cancers12123700
   IF: 6.639

List of other publications

3. Farkas, A., **Mészáros, B.**, Szarka, M., Szigeti, M., Kappelmayer, J., Szabó, M., Csánky, E., Guttman, A.: Modeling of the Desialylated Human Serum N-glycome for Molecular Diagnostic Applications in Inflammatory and Malignant Lung Diseases.
   Curr. Mol. Med. 20 (10), 765-772, 2020.
   DOI: http://dx.doi.org/10.2174/1566524020666200422085316
   IF: 2.222

4. **Mészáros, B.**, Kovács, Z., Gebri, E. Z., Jankovics, H., Vonderviszt, F., Kiss, A., Simon, Á., Botka, S., Hortobágyi, T., Guttman, A.: N-glycomic analysis of Z(IgA1) partitioned serum and salivary immunoglobulin A by capillary electrophoresis.
   Curr. Mol. Med. 20 (10), 781-788, 2020.
   DOI: http://dx.doi.org/10.2174/1566524020666200413114151
   IF: 2.222



UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

5. Gebri, E. Z., Kovács, Z., **Mészáros, B.**, Tóth, F., Simon, Á., Jankovics, H., Vonderviszt, F., Kiss, A., Guttman, A., Hortobágyi, T.: N-Glycosylation Alteration of Serum and Salivary Immunoglobulin A Is a Possible Biomarker in Oral Mucositis.
J Clin Med. 9 (6), 1-14, 2020.
DOI: http://dx.doi.org/10.3390/jcm9061747
IF: 4.241

6. **Mészáros, B.**, Járvás, G., Hajba, L., Szigeti, M., Dallos, A., Guttman, A.: Quantitative characterization of plasma treated PDMS microfluidic substrates by inverse gas chromatography.
Sens. Actuators B. Chem. 258, 1184-1190, 2018.
DOI: http://dx.doi.org/10.1016/j.snb.2017.11.185
IF: 6.393

Total IF of journals (all publications): 24,922
Total IF of journals (publications related to the dissertation): 9,844

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

19 January, 2023

DEBRECENI EGYETEM * Egyetemi és Nemzeti Könyvtár *