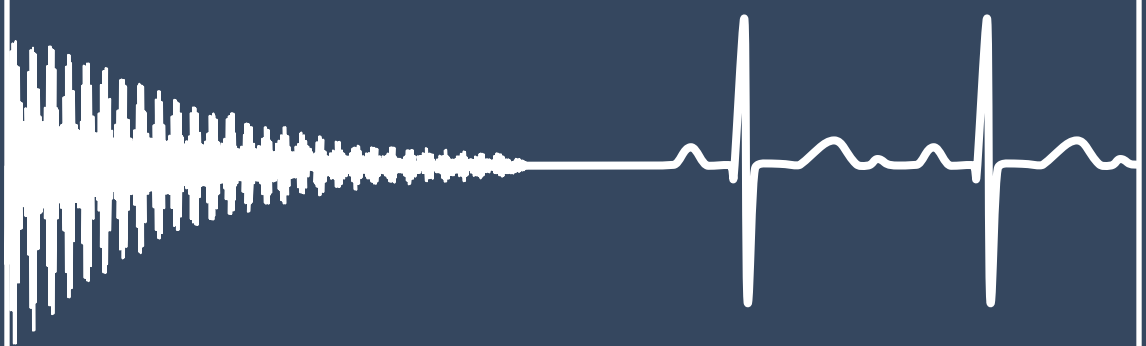
Glycosylation profiling with mass spectrometry: method development and application to cancer biomarker studies

Vreeker, G.C.M.

GLYCOSYLATION PROFILING WITH MASS SPECTROMETRY

method development and application
to cancer biomarker studies

Geertruida Cornelia Maria Vreeker



Stellingen behorende bij het proefschrift

GLYCOSYLATION PROFILING WITH MASS SPECTROMETRY

Geertruida Cornelia Maria Vreeker

1. Glycosylation profiling with MALDI-MS is a powerful method to obtain a global overview of changes in disease, but for biomarker development a more in-depth technique is probably required. *(this thesis)*

2. Glycans derived from biological material should not be evaluated individually, but always as a part of the total complex. *(this thesis)*

3. Finding a difference in serum protein glycosylation between cases and controls does not necessarily mean that this can be called a clinical biomarker, as many glycosylation changes are observed in other cancers or inflammatory diseases, making it highly unspecific. *(this thesis)*

4. The heterogeneity of breast cancer forms a challenge in finding a unifying marker for this disease. *(this thesis)*

5. To improve the understanding of cancer glycosylation, attention should not only be paid to the classification of case and control samples, but also to biological pathways connecting the glycan signatures with the tumor.

6. Collaboration between researchers from different life-science disciplines is essential and should be stimulated; the human body is one complex and the research on it should therefore be as one too.

7. Human diversity is a positive thing for evolutionary purposes, but can cause serious difficulties when it comes to biomarker discovery and disease treatment.

8. Science can never be considered objective, as the bias and prejudice of a paradigm are always present.

9. The prospect of something being hard is not a good reason not to try.

10. The ability of a child to be amazed by the smallest things in life should be cherished for as long as possible, as education in later life can disfigure this perfect happiness resulting from unknowingness.

11. Albert Einstein wrote (Tokyo, 1922): "A calm and modest life brings more happiness than the pursuit of success combined with constant restlessness." The trend and pressure of having to strive for more success and prestige all your life is one of the most toxic believes in current society.

12. The most important thing in life is not happiness, it is hope. Without hope humanity is lost.

13. De aanspreekvormen tijdens de verdediging zouden in beide richtingen toegepast moeten worden: zoals een kandidaat de commissie als "hooggeleerde/zeergeleerde opponens" aanspreekt, zou de kandidaat met een Master diploma ook als "weledelgeleerde kandidaat" aangesproken moeten worden.

GLYCOSYLATION PROFILING WITH MASS SPECTROMETRY

method development and application to cancer biomarker studies

Geertruida Cornelia Maria Vreeker

ISBN: 978-94-6402-083-0

The work presented in this thesis was performed at the Center for Proteomics and Metabolomics and Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands.
This work was supported by the society "Genootschap ter ondersteuning van de vroege opsporing van kanker" (Lisse, The Netherlands) to further endorse the development of a blood-based test for early detection of cancer (no grant number applicable).
Cover design & layout: Geertruida Cornelia Maria Vreeker
Printed by: Gildeprint – The Netherlands

GLYCOSYLATION PROFILING WITH MASS SPECTROMETRY

method development and application to cancer biomarker studies

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. dr. ir. H. Bijl,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 28 oktober 2021
klokke 10:00 uur

door

Geertruida Cornelia Maria Vreeker
geboren te Hoorn
op 20 juli 1992

Promotores

Prof. dr. M. Wuhrer

Prof. dr. R.A.E.M. Tollenaar

Co-promotor

Dr. W.E. Mesker

Leden Promotie Commissie

Prof. dr. C.M. Cobbaert

Prof. dr. C.H. Hokke

Prof. dr. D.J. Lefeber

Department of Neurology, Faculty of Medical Sciences,
Radboud University/Radboud UMC, Nijmegen, The Netherlands

Dr. T.M. Luider

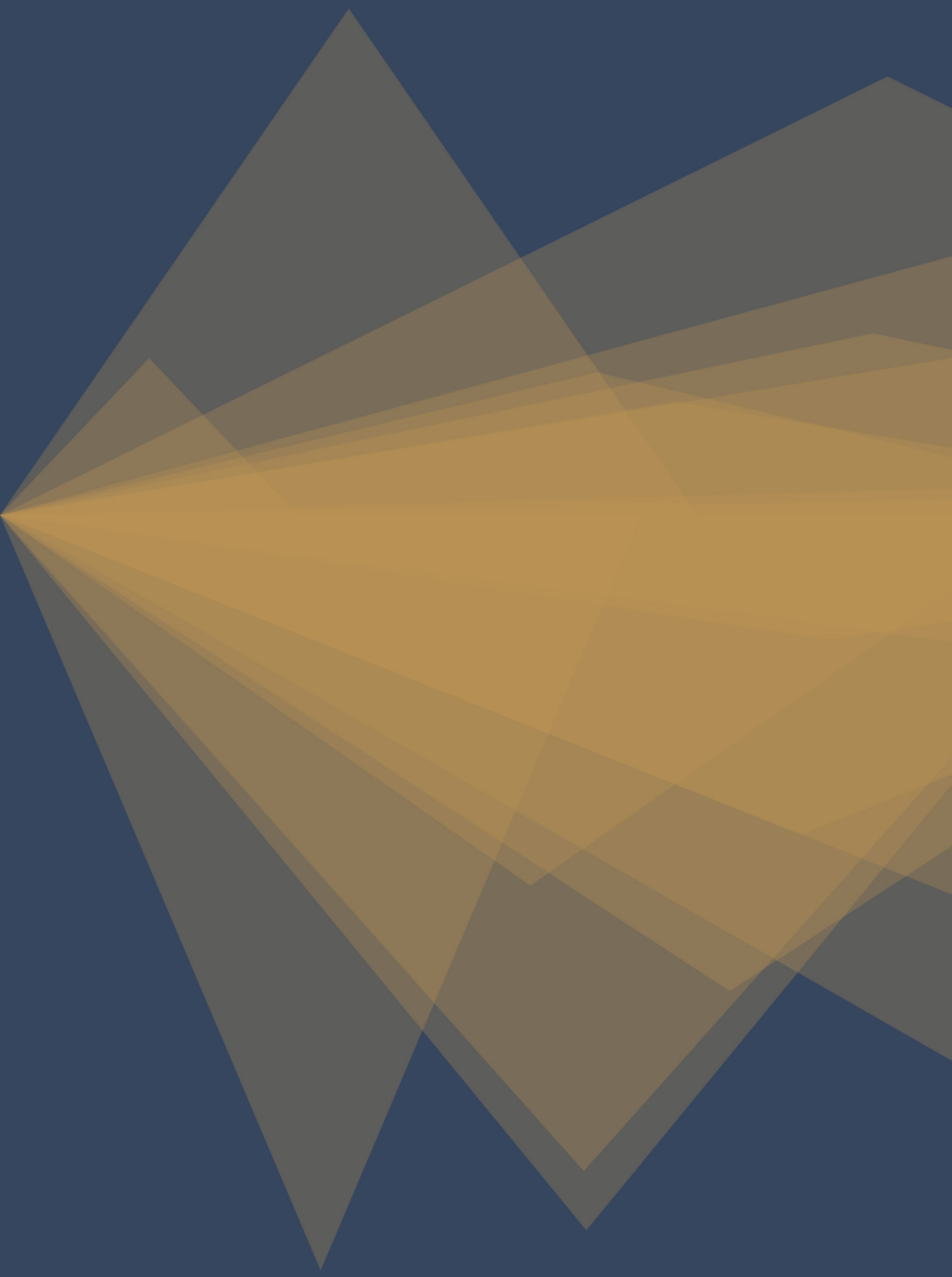
Department of Neurology, Erasmus MC, Rotterdam, The Netherlands

"You are braver than you believe, stronger than you seem, and smarter than you think" - *Christopher Robin*

— A. A. Milne, Winnie-the-Pooh

To all persons that had, have or will have to face cancer and to their loved ones

Table of contents



CHAPTER 1

Introduction

General introduction

Cancer is one of the leading causes of death worldwide. Breast cancer is the most common cancer in females. Pancreatic cancer has a very poor prognosis, with almost as many deaths as incidences.[1] The detection of cancer at early stage is challenging, but highly important as this often comes with a better prognosis and more treatment options. For breast cancer population screening and pancreatic cancer high-risk screening imaging techniques are used for detection. As is true for almost any diagnostic test, current screening methods are not optimal in terms of sensitivity and specificity, resulting in false positives and false negatives, which could have far-reaching consequences for patients. The need for new biomarkers and the development of new highly sensitive and specific methods is high, without compromising on low invasiveness and risk for the patients. To this end blood-based screening has great potential and as a result multiple blood based markers have been reported from omics-discovery studies, although none of these have provided the desired sensitivity and specificity that are needed for clinical implementation for population/high-risk screening.[2] In addition, blood-based biomarkers could add to the process of clinical decision making to improve treatment. In cancer, the expression of glycosyltransferases and glycosidases is altered, which leads to changes in glycosylation of proteins and lipids.[3,4] Qualitative and quantitative analysis of the glycosylation of proteins might provide a better understanding of these changes. Moreover, recent advances in glycoanalytical strategies have allowed glyco(proteo)mics to be implemented in cancer biomarker studies for exploring glycans as a potential source of biomarkers.

The research described in this thesis was focused on development of released *N*-glycosylation analysis methods for dried blood spots as well as for serum, of which the latter was applied to a pancreatic cancer case-control study. Furthermore a breast cancer case-control study was performed. In this introduction a clinical background of breast- and pancreatic cancer is provided, the basics of protein glycosylation are described and the analysis of glycosylation from blood-derived samples with mass spectrometry is explained.

Cancer

For 2018 it was estimated that 18.1 million new cancer cases would be diagnosed worldwide and that 9.6 million people would die from the disease.[1] Of these new cancer cases it was estimated that breast cancer would account for over 2 million incidences and 626 thousand deaths. For pancreatic cancer the number of incidences was estimated to be much lower (459 thousand), however, because of its poor prognosis the number of deaths (432 thousand) is almost equal to the number of incidences.[1] It is predicted that the number of pancreatic cancer related deaths will in the future be even higher than the deaths related to breast cancer, while the incidence of breast cancer is more than four times higher.[1] As cancer is the first or second cause of death in 91 countries, the number of studies related to cancer is vast.

Much attention is paid to study the biology of cancer and its pathophysiology. Cancer cells often develop from normal tissues, caused by mutated genes. These genes can be inherited, and thus be present at birth, but often these are acquired through exposure by environmental factors (e.g. radiation, chemicals, asbestos, bacteria/viruses) or lifestyle-related factors (e.g. cigarette smoke, alcohol, sunlight, food-related factors). The result of the mutated genes is not presented from the onset. It takes time for the cancer cells to develop and even more time before symptoms will appear. Some cancers, such as pancreatic cancer, are often diagnosed when these have already reached an advanced stage, decreasing the possibility for curative treatment and patient survival chances.[5]

Pancreatic cancer

Pancreatic cancer is the seventh cause of cancer deaths, which is not due to a high incidence, but rather a result of aggressiveness of the cancer, leading to a poor prognosis for pancreatic cancer patients.[1] The pancreas is located in the abdomen, where it is surrounded by other organs and large blood vessels (i.e. aorta and inferior vena cava), the pancreas is therefore a relatively inaccessible organ.[6] It contains a duct system through which produced enzymes are transported to the digestive system. These digestive enzymes are produced in lobules located throughout the pancreas being part of the exocrine system.[6] The pancreas also has an endocrine function: it produces hormones to regulate the blood sugar level.[7] Most pancreatic cancers (85-90%) are pancreatic ductal adenocarcinomas (PDAC) which are originating from the exocrine system. Neuroendocrine tumors (NET) are only found in 1-2% of the pancreatic cancer patients.[5]

In most patients that are diagnosed with pancreatic cancer the tumor is at an advanced stage and resection is in many cases not possible. Pancreatic cancer is at an advanced stage when the celiac axis or superior mesenteric artery are affected

or in case of metastasis.[8] The symptoms of pancreatic cancer are often aspecific, including malaise, weight loss and pain in the abdomen. Most pancreatic cancers have metastasized ahead of manifestation of the disease and in cases the initial diagnosis was a localized cancer, the patients still die from metastatic or recurrent cancer.[9] Notably, patients are also diagnosed with pancreatic cancer as a side-finding in scans for unrelated matters (e.g. trauma).[9] To increase the percentage of patients where resection is still possible, screening modalities for genetic high risk groups are being developed to diagnose pancreatic cancer at an earlier stage.[10]

As the incidence of pancreatic cancer is relatively low, the benefit of a general population screening does not outweigh the downsides such as false-positives and high costs. Therefore currently only persons at high risk are tested. Besides genetic risk factors, also clinical risk factors for pancreatic cancer were found: chronic pancreatitis, diabetes and a family history of pancreatic cancer, but also age, smoking and obesity, which are nonspecific.[9]

Some serum biomarkers are available for pancreatic cancer, such as CA19-9. This marker has a low sensitivity and specificity which prevents its use in screening. Instead, CA19-9 is used for monitoring of patients during treatment.[9] Promising results have been reported from protein biomarker discovery studies on pancreatic juice, but this liquid is difficult to obtain with inherent risks for the patient.[9,11] Alternatively, imaging techniques are used for screening, however with ultrasonography the sensitivity is relatively low because of the location of the pancreas in the abdomen.[12] Contrast-enhanced computerized tomography is also used for diagnostic purposes and can show relatively high sensitivity and specificity, but this technique is more invasive.[9] A third imaging technique that is used for pancreatic cancer diagnosis is magnetic resonance imaging (MRI), but this technique is expensive and less convenient for the patients as they have to remain motionless during the scan.[13] Nevertheless, annual MRI screening of high risk patients was shown to be successful to detect pancreatic cancer at a resectable stage.[14]

The development of a screening test for pancreatic cancer is challenging, but not limited to the above mentioned difficulties. False-positive cases are highly undesired in pancreatic cancer as the follow-up evaluations are relatively invasive[9], besides the influence on the mental wellbeing of the patient. Moreover, if a screening test would diagnose patients at an earlier stage, it does not necessarily mean that the natural development of the cancer can be altered in order to increase survival time.[9] This should be taken into account when evaluating a screening method, as implementing a screening test is only worth it when it improves the prognosis or well-being of the patient. For pancreatic cancer the higher resection rate and better survival of screened patients (often diagnosed at earlier stage) compared to sporadic pancreatic cancer cases suggests that it meets the requirement.[14]

Breast cancer

Breast cancer is the leading cancer diagnosed in women. Of all cancer incidences in women, more than 24% is breast cancer. Cancer mortality in women is in 15% of the cases the result of breast cancer.[1] It is predicted that within thirty years breast cancer incidence will reach more than 3 million new cases annually due to population aging and westernization of the developing countries.[15]

The breast is located on the outside of the chest muscles, under the skin. The nipple is in the center of the skin and is under the skin surrounded by 15-20 lobes. These lobes consist of 20-40 lobules, which produce the milk that is in the end excreted by the nipple. The breast also consists of fatty tissue, of which the amount determines the size of the breast. Additionally, blood vessels and connective tissue are present.[16,17] Most breast cancers start in the milk-producing and -transport system of the breast (e.g. the lobules, milk-ducts), but can be caused by the expression of multiple different genes.[18] Breast cancer is genetically and clinically very heterogeneous, and multiple different classification systems exist (e.g. receptor status, DNA classification).[19] In addition, prognostic staging (TNM classification) of breast cancer is based on tumor size (T), lymph nodes involvement (N) and metastasis (M) of the disease.

Population-based breast cancer screening programs are carried out to reduce mortality, and for example in the Netherlands consist of mammography every two years for women aged between 50 and 75 years.[20] This technique consists of X-ray photos taken from the breasts in search for abnormalities. Unfortunately, the percentage of cases that were not detected by mammography lies around 25%. In addition, around 75% of the women that are send to the hospital for additional tests, turn out to be a false-positive of the mammography test.[21] From these results it is no surprise that there is an ongoing debate on the use of mammography for breast cancer screening.[22] Although there is clear evidence that the use of mammography for screening of breast cancer reduces mortality,[21,22] there are also multiple downsides to this method. To begin with the false-positive results, which lead to unnecessary biopsies and also feelings of anxiety and stress for the patients. Additionally the risk of overdiagnosis is high, resulting in demanding treatments for patients where the cancer will not result in death.[22]

For breast cancer, risk factors are genetic mutations (e.g. BRCA1, BRCA2) and a family history with breast cancer, but amongst others also menarche at younger age, late first pregnancy (>30 years) or nulliparity, late menopause, oral contraceptives and a history of chest irradiation increase the risk. The average life-time risk for breast cancer is less than 15%, however, the above mentioned factors can increase this number for the individual patients. The BRCA1 and BRCA2 mutations increase the lifetime risk (LTR) of breast cancer up to 82%, however only

15% of the familial breast cancers is explained by these genes.[23] Other known, but rare, mutations connected to increased breast cancer risk are PTEN (LTR 85%), TP53 (LTR 25%), CDH1 (LTR 39%) and STK11 (LTR 32%).[23] Additionally, multiple DNA repair genes interacting with BRCA mutations or -pathways which double the risk of breast cancer have been identified (CHEK2, BRIP1, ATM, PALB2).[23]

Breast cancer testing in women carrying the BRCA genes starts at younger age, as these mutations can lead to the develop of cancers at younger age. However, mammography has some disadvantages in these cases, as younger women often have denser breast tissue which negatively influences the sensitivity of mammography. Another technique which is often used for the detection of breast cancer is MRI. Compared to mammography, MRI is relatively expensive. However, MRI has a high sensitivity in breast cancer, although the specificity is relatively low (more false positives). Notably, for young high risk patients the combination of mammography and MRI increased sensitivity.[24]

As breast cancer is a very heterogeneous disease[19] and screening challenges such as overdiagnosis and interval breast cancer (developing within 12 months after a mammography test considered normal) are present, additional research on screening methods and early cancer detection are urgently needed. New high sensitivity and specificity tests need to be developed to be able to detect all different breast cancer subtypes and to prevent overdiagnosis.

Glycosylation

The human genome consists of approximately 30,000 genes that each code for a specific protein. For each gene various molecular protein species are found, which are called proteoforms[25], resulting in more than one million different proteins presented in the human body.[26] Various types of proteoforms result from splicing upon human deoxyribonucleic acid (DNA) transcription, where introns are removed from the messenger ribonucleic acid (RNA), or from post-translational modifications (PTMs), in which structural changes occur after translation of the RNA to proteins.[27] These PTMs have influence on for example the structure and solubility, leading to changes in function and activity of the proteins.[28] Many different types of PTMs are known, all with various functions and each with a very specific presence and abundance on proteins. Most of these modifications are enzyme-catalyzed and several are reversible.[27,28]

A frequently observed PTM is glycosylation, which is defined as "the enzyme-catalyzed covalent attachment of carbohydrate to a polypeptide, lipid, polynucleotide, carbohydrate, or other organic compound, generally catalyzed by glycosyltransferases, utilizing specific sugar nucleotide donor substrates".[29] Glycosylation is commonly found on secretory as well as membrane proteins in single cell- and multicellular organisms.

Monosaccharides

Glycans are built from monosaccharides, which in solution exist in an equilibrium of the cyclic and linear form. Upon cyclization of a linear monosaccharide an additional chiral center is formed at C-1, resulting in a so-called α- and β-anomer (exemplified for glucose in Figure 1). The anomeric center of a monosaccharide can interconvert between both isomers via the linear structure. However, in a oligo- or polysaccharide, only one monosaccharide can interconvert at the anomic center, namely the one that is not bound to another monosaccharide and this is called the reducing end. This name originates from the redox reaction that may be performed on the aldehyde in the cyclization (Figure 1) resulting in an oxidized carbonyl group.

Glycans are complex oligosaccharides which, in contrast to linear polysaccharides such as hyaluronan and heparin, often are branched structures. This branching increases the complexity of the glycans and results in many isomeric structures, which subsequently might influence the function. Moreover, the variety of monosaccharide building blocks that are commonly present in glycan structures further increases complexity. In addition, monosaccharides are linked via glycosidic bonds, which in theory can be formed between all hydroxyl-groups present on the monosaccharides.

The great variety of possible glycans makes the study of their function challenging. Glycans are influencing the structure and conformation of proteins, which leads to changes in protein interactions and protein activity.[30,31] In addition, glycans are involved in many biological processes, such as cell recognition, cell signaling, inflammation, immune defense and regulation processes including cell survival and growth.[32]

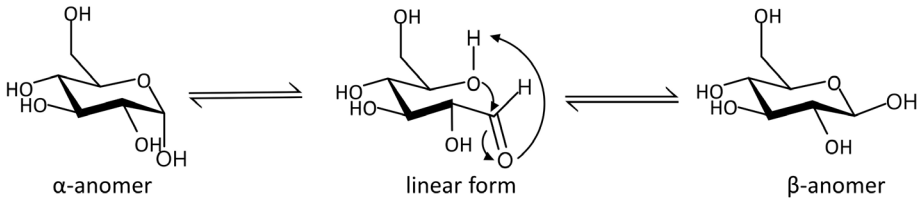


Figure 1. Equilibrium of cyclic and linear glucose.

Protein glycosylation: structures and biosynthesis

Glycosylation is found on many different compounds in human cells, for example on proteins. The most common glycan structures found on proteins are *N*- and *O*-glycans. Enzymes couple monosaccharides to form branched and often complex polysaccharides. The substrate specificity of these glycosyltransferases determines the primary structure of a glycan.[33] *N*-linked glycans are linked to an asparagine (Asn) via the nitrogen atom from the Asn amine group. Not all asparagines in a protein are potential glycosylation sites. The specific amino acid sequence Asn-X-Ser/Thr, where X cannot be a proline, is referred to as a structural motif for potential *N*-glycan attachment.

N-glycans all share the same core structure consisting of three mannoses (Man) and two *N*-acetylglucosamines (GlcNAc): Manα1–6(Manα1–3)Manβ1–4GlcNAcβ1–4GlcNAcβ1-Asn-X-Ser/Thr. Three different classes of *N*-glycans are defined: oligomannose-type, complex-type and hybrid-type *N*-glycans. The oligomannose glycans consist of mannoses only in their core structure; complex glycans have on both outer mannoses of the core so-called 'antennae' attached, which start with GlcNAc and are followed by a variety of different monosaccharides; the hybrid type *N*-glycans contain on the Manα1-6 arm only mannoses, while on the Manα1-3 arm a complex antenna is attached (Figure 2a). The most common monosaccharides found in human *N*-glycans are besides GlcNAc and Man: fucose (Fuc), galactose (Gal) and *N*-acetylneuraminic acid (Neu5Ac), the latter is also referred to as "sialic acid".

Roughly, the biosynthesis of *N*-glycans is a two-step process, which takes place in the endoplasmic reticulum (ER) and the Golgi. The build-up of the glycan starts in the ER with a dolichol phosphate to which monosaccharides are one by one

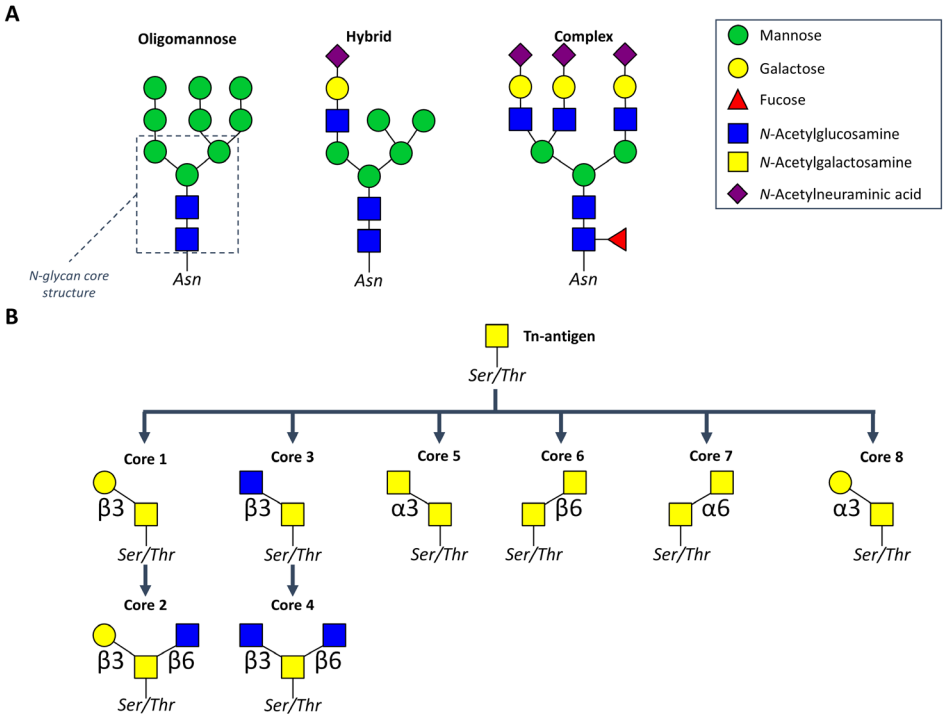


Figure 2. (a) *N*-glycan types (oligomannose, hybrid and complex) and general *N*-glycan core. (b) Tn-antigen and *O*-glycan cores.

coupled by enzymes. An oligomannose glycan, consisting of nine mannoses and extended with three glucoses on the α1-3 arm, is formed. This glycan is in its entirety transferred to the asparagine of the glycosylation site of the protein and all other glycans are derived from this single glycan structure. Subsequently the glycan is trimmed and rebuilt in the Golgi, which gives the protein its specific function.[34] Although this pathway might seem a devious process, the glycan intermediates serve as a quality control for the protein.[34]

The other most commonly found glycan structures on proteins are *N*-acetylgalactosamine (GalNAc) *O*-glycans. These glycans are linked to Ser or Thr via an oxygen atom (GalNAcα-Ser/Thr). Unlike *N*-glycans, *O*-glycans do not share a single core structure: eight different core structures are known, of which core 1 and 2 are most often found (Figure 2b).[35,36] All *O*-glycan cores contain an *N*-acetyl galactosamine and besides this monosaccharide *O*-glycans also can contain galactose, *N*-acetyl glucosamine, fucose and sialic acid sugars.[35] In contrast to *N*-glycans, *O*-glycans are directly synthesized onto the protein (per monosaccharide) after *N*-glycosylation, folding and protein oligomerization.[36]

Glycosylation and cancer

In tumor cells alterations in glycosylation are often observed and it has been suggested that these play specific roles in the progression of the disease.[3,37,38] In cells, major changes of *N*- and *O*- glycosylation are associated with transformation of the cell in early stage as well as in progression of the disease and metastasis.[39] Furthermore, an increase in specific glycosyltransferases in cancer influences the production of glycans and alters the 'normal' glycosylation, which subsequently may influence the structure and function of the glycoprotein.[39] The conformation of the glycoproteins, as well as the interaction with other molecules and oligomerization can be changed, which may contribute to the progression of the tumor. Also, the interaction between the tumor and its microenvironment might be altered.[3,39,40] On tumor cells specific glycan structure are often found to be highly expressed, of which the Tn-antigen (Figure 2b) is an example.[39]

Altered *N*-glycosylation features which are often reported are increased branching and fucosylation (upregulation of FuT8) in cancer patients.[41] The latter has been associated with the regulation of cell proliferation[39], but also decreased cell adhesion and cancer development.[3,42] The increased branching of glycans in cancer cells has been associated with the upregulation of MGAT5 and is involved in tumor growth and metastasis.[42–45] Additionally, expression of sialyl Lewis antigens (sLe) is upregulated.[39,41] The glycosylation epitope sLeX is found on the cell surface, but also on glycoproteins in serum of cancer patients.[46] Notably, it was reported that the most lethal cancers, which are pancreatic, lung and gastric cancer, show the highest serum sLeX levels.[47]

Analytical methodology

To understand the role of glycosylation in cancer and other diseases, information on glycosylation changes, in terms of structures and/or amounts present, is collected. This data can be gathered from different human sample types, such as whole blood samples, serum/plasma samples, tissue, saliva and urine, depending on the research question and goal.[48–52]

Glycosylation analysis can be performed on different levels. It can be analyzed from intact glycoproteins, from glycopeptides or from solely the glycans themselves. Each of the levels provide different information on the glycosylation. With intact analysis all modification on a glycoprotein are presented in the system in which they function. The number of PTM's per protein can be analyzed and also the information on combinations of PTM's can provide insights in the function of the protein. However, the analysis of intact proteins can be challenging as a good separation of all different proteoforms is difficult to establish, particularly when proteins contain many glycosylation sites or are part of a complex mixture. Alternatively glycopeptide analysis can link PTM's to the exact location on the protein, giving insight in the occupation and type of glycans on a specific glycosylation site. This however is still challenging with complex mixtures of many different glycoproteins.

The focus of this thesis is on the analysis of released glycans from glycoproteins in complex samples (e.g. serum and blood). Although the obtained glycan information cannot be linked to glycosylation sites or specific proteins on this level of analysis, highly informative data on overall glycan heterogeneity can be obtained from very complex samples. Changes in certain types of glycosylation, such as levels of different antennarities, fucosylation or sialylation, might give an indication on processes that are related to the disease investigated. Moreover, these up- and down-regulation of certain types of glycans might still be linked to certain glycoproteins, as information on the glycosylation of these specific proteins is available.[53]

Over the years, various analytical methods have been described for the structural, qualitative or (semi-)quantitative analysis of glycans. Nuclear magnetic resonance (NMR) has for example been used to analyze the structure of glycans.[54–56] The advantage of NMR is that the exact structure of the glycan can be determined and thus can differentiate between isomers. However, the sample needs to be nearly pure, because mixtures of glycans can cause great difficulties in interpreting the data. In addition, the interpretation of NMR spectra from large structures is also challenging, due to the similar environments many of the protons encounter. Another limiting aspect of NMR is the amount of sample required for analysis, which is extremely

high in comparison with a sensitive technique such as fluorescence detection, UV detection or mass spectrometry (MS).[57]

Fluorescence detection and UV detection are often used after liquid chromatography (LC) separation of labeled glycans. Also electrospray ionization (ESI)-MS is generally used after LC separation of glycans with and without prior labeling. LC separations of glycans have been extensively reviewed for various stationary phases, specifically hydrophilic interaction liquid chromatography (HILIC), porous graphitized carbon (PGC) chromatography and reversed phase (RP) chromatography.[58–62] The different stationary phases are used for different research questions and purposes. HILIC chromatography is often used for complex samples as it has a high peak capacity. PGC however, has the great advantage of separating glycan isomers, including linkage isomers. On the other hand, RP-LC is usually highly reproducible and also widely available among research laboratories.[62,63] Additionally, glycan separation with capillary electrophoresis (CE) is also a possibility and is often performed in combination with fluorescence or MS detection.[64–7]2 A more in-depth analysis of glycan structures can be done using endo-glycosidases. A sequential use of these enzymes in combination with the techniques described above, can reveal the monosaccharide structure of glycans.

In this thesis matrix-assisted laser desorption/ionization (MALDI)-MS is used as an analytical technique for released glycans. Both MALDI-time of flight (TOF)-MS and MALDI-Fourier-transform ion cyclotron resonance (FTICR)-MS are evaluated and applied to released glycan analysis.

Serum and blood sampling

In the research described this thesis, glycosylation analysis is performed on serum, plasma, whole blood and dried blood. In general, the quality of preanalytical processing of blood samples is extremely important for the quality of the outcomes of the research. Inaccurate results might be obtained when the preanalytical processing is not thoroughly thought out prior to cohort collection. Technical variables such as variation in collection tubes, sampling protocols and storage temperature might lead to biases in the outcome of the research, but also biological variables such as age, gender, ethnicity and lifestyle factors might play a significant role in this.[73] For proteomics studies it was shown that the standardization of preanalytical processing is of great importance[73], while it appears that glycans are less influenced by the preanalytical processing conditions.[74]

Research has indicated that glycosylation profiles are very stable compared to other serum or plasma analytes such as proteins.[74] For plasma samples no effect of storage conditions before centrifugation as well as longterm storage at -20 versus -80 °C has been found on MALDI-TOF-MS *N*-glycosylation profiles. Also using different

collection tube additives and overnight shaking of the sample did not introduce any difference in the profiles.[74]

The stability of glycans in DBS is investigated in this thesis. In the past stability of glycans in DBS has already been evaluated at room temperature and -80 °C, concluding that also in DBS the glycome is stable.[49] Proteins and peptides however, degrade relatively fast in DBS.[75] The stability of the glycome in DBS makes the use of DBS samples for glycan analysis very attractive for sample collection. Especially because it is less invasive compared to venipuncture and might even be performed by patients themselves.

Glycan release

Intact glycan analysis is an upcoming field, but before that glycosylation was mainly analyzed on released glycan level.[76] Released glycan analysis is still widely applied and also in this thesis this approach is used. An important criterium for a release method is that it should not have a bias towards certain glycans, and thus should only be selective for the cleavage site. In addition, it should not modify the glycan except for the cleavage of the target bond. Moreover, the reaction has to be efficient enough that a sufficient amount of released glycan material can be recovered. It should also be mentioned that it is highly desirable that a free reducing end is available for labeling after glycan release.[76]

Both chemical and enzymatic releases of glycans are possible. Chemical release of glycans is performed via for example β-elimination at alkaline conditions. With this method *N*-glycans as well as *O*-glycans can be released, dependent on the alkaline conditions used: with harsh alkaline conditions the *N*-glycans are released from the proteins, whereas with dilute alkaline conditions only *O*-glycans are released. Notably, with this reaction the protein backbone is heavily affected and the GlcNAcs are deacetylated.[77]

N-glycans are often cleaved using peptide-*N*-glycosidase F (PNGase F). This enzyme breaks the bond between the Asn and GlcNAc, which results in an aspartic acid residue and a glycosylamine.[78] The stability of the amine is not very high at a neutral pH, resulting in a hydrolysis of the amine to a hydroxide group. The reaction takes place under mild conditions, leaving the protein backbone intact. In addition, it is specific for *N*-glycans, as PNGase F only cleaves glycans that are attached to an asparagine.[79] Unfortunately no enzyme has been found yet that cleaves all types of *O*-glycans.[78]

Glycan derivatization and purification

Derivatization reactions are performed to improve the molecular properties of an analyte for analysis. On glycans, often derivatized sites are the reducing end,

the carboxylic acid of the sialic acids and all hydroxyl groups in the molecule. Reducing end modifications are usually labels which can be used to analyze glycans with fluorescence or UV detection.[80–83] Also for MS detection labels are used, which might enhance sensitivity or provide semi-quantitative information upon fragmentation.[84–86] Additionally, stable isotope labelling can be used for the direct comparison of samples or as an internal standard.[87–89] In this thesis the reducing end of the glycans is not modified other than the reduction of the reducing end in the case of the *O*-glycans which happens during the release reaction.

Sialic acid modifications are essential for the analysis of sialylated glycans with MALDI-MS. In-source/post-source fragmentation, resulting in the loss of sialic acids, is often observed with MALDI-MS.[90,91] Previous studies have reported that this is due to the carboxylic acid that is located next to the glycosidic bond. With an internal reaction the glycosidic bond is broken, resulting in the loss of the complete sialic acid.[90,92] This type of cleavage has also been observed for peptides.[93,94] To stabilize the sialic acids they can be methyl- or ethyl esterified.[50,90,91,95,96] In addition the carboxylic acid can be amidated and/or permethylated.[97,98]

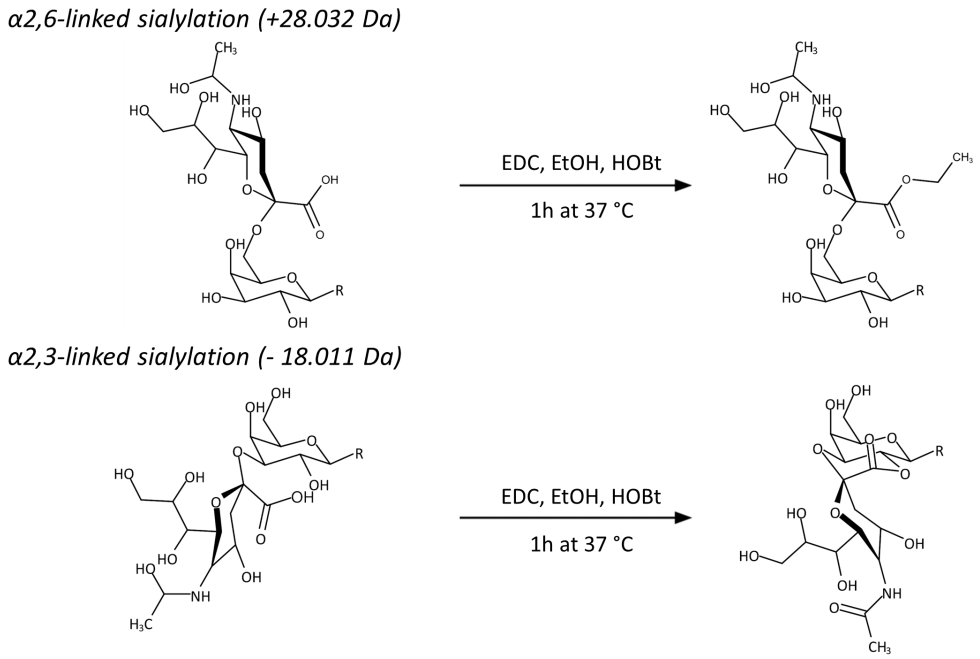


Figure 3. Ethyl esterification reaction of α2,6-linked sialic acids, and lactonization reaction of α2,3-linked sialic acids.

In this thesis ethyl esterification (*N*-glycans) and permethylation (*O*-glycans) are performed to stabilize the sialic acids. It should be noted that many of the above

mentioned derivatization reactions are linkage-specific. Sialic acids are linked to their neighboring galactose (in the case of *N*-glycans) with an α2,3- or α2,6-linkage. Conformationally the carboxylic acid of the α2,3-linked sialic acid is in proximity in space to the galactose, which leads to an internal reaction forming a lactonized structure (Figure 3). On the contrary, the carboxylic acid of the α2,6-linked sialic acid has a significant distance in space from the galactose and forms an ester at the carboxylic acid instead of the internal ring (Figure 3). In the case of ethyl esterification, the mass differences obtained in comparison with nonderivatized are -18.011 Da and +28.031 Da for the α2,3- and α2,6-linked sialic acids respectively. In MALDI-MS analysis this mass difference between the glycans containing α2,3- or α2,6-linked sialic acids, can be used to gather linkage-specific information on the glycosylation in a sample.[91]

In permethylation of the *O*-glycans not only the carboxylic acid is modified but all hydrogens of the hydroxyl and amine groups are replaced by methyl groups. This modification decreases the hydrophilicity of the molecule, which is beneficial in for example RP separations, but also for sensitivity in ESI-MS.[62,99,100]

MALDI-MS

Biomarker research is often performed on large sample cohorts. To minimize batch effects, it is preferred that samples are measured in a short time period. A high-throughput analysis method such as MALDI-MS can fulfill this need as the general measurement time per sample is 10 to 20 seconds, depending on laser frequency and analyte concentration.

MALDI is a soft ionization method which was introduced in 1985. It is called 'soft' as MALDI ionizes analytes without breaking the chemical bonds in the analyte molecule.[101] This allows direct mass measurement of biomolecules, such as glycans. Nevertheless, for certain glycan structures in-source fragmentation occurs in small amounts upon MALDI ionization.[92,102] For example, the loss of sialic acids as mentioned earlier in this Introduction and consequently this requires a different strategy (such as derivatization).

In general, MALDI analysis starts with applying a small drop (few microliters) of the sample onto a MALDI target together with a matrix. The matrix consists of small organic molecules which have a strong absorption at the wavelength of the laser.[103] The sample and matrix then dry together to form analyte containing matrix crystals on the MALDI target plate. Under vacuum conditions in the source these crystals are ablated with shots of an ultraviolet laser, resulting in analyte ions. The ions are then transferred to the mass analyzer, which can be a TOF or FTICR instrument, and mass-to-charge (m/z) ratios are determined.[104]

There are also some challenges generally encountered in MALDI-MS. First,

the dried sample/matrix spots on the MALDI target are often not homogeneous. For glycan analysis generally 2,5-dihydrobenzoic acid (DHB) matrix is used, which forms various large crystals thus resulting in heterogeneous spots. A rather large shot-to-shot variation is therefore frequently observed in MALDI-MS analysis.[105] To compensate for this effect and to improve repeatability of measurements multiple shots on a spot can be averaged. When these are recorded in an unbiased random pattern the average spectrum obtained from a spot is relatively stable.

Second, in MALDI-MS chemical noise is often present in the spectra. Chemical noise can arise from impurities in the sample, but more challenging is the chemical noise caused by matrix ions.[106] Especially with low analyte concentrations and in the lower mass region (below *m*/*z* 800) many signals originating from matrix clusters are visible. Analysis of small glycans at low concentrations can thus become challenging.

Mass analyzers

Ions are evacuated from the MALDI-source and accelerated by an electric field towards the mass analyzer, in TOF this is the flight tube. This tube is field-free and the ions are separated here based on their velocities. From the time it takes the ions to move through the flight tube to the detector the *m*/*z* ratios are determined. A TOF analyzer can analyze ions over a very broad mass range, making it very suitable to combine with a MALDI-source. These measurements can be performed in linear mode, as well as reflectron mode. A reflectron deflects the ions back into the flight tube to a second detector. Here, the kinetic energy dispersion of ions with the same *m*/*z* is corrected, which improves mass resolution. Ions with a higher kinetic energy will travel deeper into the reflectron, giving them a slightly longer path than ions with a lower kinetic energy, which results in the ions reaching the detector at the same time. However, using reflectron mode results in a mass range limitation and a loss of sensitivity.[103]

Another analyzer which can be combined with a MALDI-source is FTICR. The ions are transported to an ion trap cell within a magnetic field. For detection purposes, the ions are excited and the corresponding cyclotron frequencies are measured from an induced image current. In Fourier transform MS all ions are excited simultaneously, resulting in a complex wave, which can be transformed into individual frequencies. The obtained frequency spectra is Fourier transformed into a mass spectrum that displays *m*/*z*-values for each species with a corresponding intensity.[103]

Scope

The aim of this thesis is to develop robust and high-throughput methods for the analysis of *N*-glycosylation from serum proteins, to facilitate biomarker discovery focused at early detection of cancer. For this purpose a glycan profiling method with a MALDI-FTICR-MS platform was established and applied to a pancreatic cancer cohort.

In Chapter 1 of this thesis an overview was provided on cancer, with the emphasis on pancreatic and breast cancer from a clinical point of view. In addition a general introduction to glycosylation and glycan analysis was given.

The second part of this thesis, consisting of Chapters 2, 3 and 4, focuses on method development for glycosylation analysis. A high-throughput method including MALDI-TOF-MS for the analysis of *N*-glycans from serum was reported earlier and this method was used as a starting point for the development of the method described in Chapter 2. Points of improvement are the prevention of side reactions and the automation of cotton-HILIC SPE. Moreover, the glycans are analyzed with ultrahigh resolution MALDI-FTICR-MS, resulting in resolved instead of overlapping peaks and in highly confident glycan assignments. In Chapter 3 the method from the previous chapter is extended to the analysis of DBS instead of serum. For clinical applications the sampling of a DBS is desired as it is less invasive and cumbersome than the collection of venous blood specimens. A method for the analysis of released glycans from DBS was developed and evaluated for different storage conditions. Chapter 4 elaborates on the MALDI-FTICR-MS measurements of glycans from complex samples and explores the use of ultrahigh-resolution approaches for the comprehensive detection of *N*- and *O*-glycan signals over a vast *m*/*z* range.

Chapter 5 describes the application of the method from Chapter 2 to a pancreatic cancer biomarker study. Here, case and control samples from two independent cohorts are compared for their *N*-glycosylation profiles. This allows to study the glycosylation differences between pancreatic cancer patients and healthy controls and might in future be used for the improved diagnosis of pancreatic cancer. In Chapter 6 a similar method is applied to a breast cancer case-control biomarker study. Here, the focus is on the comparison of the results with findings described in literature.

In Chapter 7 a general discussion is given on the work described in the previous chapters. Finally, a perspective is provided on method development and translational research.

References

(1) Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R. L.; Torre, L. A.; Jemal, A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA. Cancer J. Clin. 2018, 68 (6), 394–424. https://doi.org/10.3322/caac.21492.

(2) Parker, C. E.; Borchers, C. H. Mass Spectrometry Based Biomarker Discovery, Verification, and Validation - Quality Assurance and Control of Protein Biomarker Assays. Mol. Oncol. 2014, 8 (4), 840–858. https://doi.org/10.1016/j.molonc.2014.03.006.

(3) Vajaria, B. N.; Patel, P. S. Glycosylation: A Hallmark of Cancer? Glycoconj. J. 2017, 34 (2), 147–156. https://doi.org/10.1007/s10719-016-9755-2.

(4) Reis, C. A.; Osorio, H.; Silva, L.; Gomes, C.; David, L. Alterations in Glycosylation as Biomarkers for Cancer Detection. J. Clin. Pathol. 2010, 63 (4), 322–329. https://doi.org/10.1136/jcp.2009.071035.

(5) Haugk, B.; Raman, S. Pancreatic Pathology: An Update. Surg. (United Kingdom) 2016, 34 (6), 273–281. https://doi.org/10.1016/j.mpsur.2016.03.012.

(6) Bockman, D. E. Anatomy of the Pancreas. In The Pancreas: Biology, Pathobiology and Disease; Raven Press: New York, 1993; p 8.

(7) Kahl, S.; Malfertheiner, P. Exocrine and Endocrine Pancreatic Insufficiency after Pancreatic Surgery. Best Pract. Res. Clin. Gastroenterol. 2004, 18 (5), 947–955. https://doi.org/10.1016/S1521-6918(04)00089-7.

(8) Bilimoria, K. Y.; Bentrem, D. J.; Ko, C. Y.; Ritchey, J.; Stewart, A. K.; Winchester, D. P.; Talamonti, M. S. Validation of the 6th Edition AJCC Pancreatic Cancer Staging System: Report from the National Cancer Database. Cancer 2007, 110 (4), 738–744. https://doi.org/10.1002/cncr.22852.

(9) Poruk, K. E.; Firpo, M. A.; Adler, D. G.; Mulvihill, S. J. Screening for Pancreatic Cancer: Why, How, and Who? Ann. Surg. 2013, 257 (1), 17–26. https://doi.org/10.1097/SLA.0b013e31825ffbfb.

(10) Singhi, A. D.; Koay, E. J.; Chari, S. T.; Maitra, A. Early Detection of Pancreatic Cancer: Opportunities and Challenges. Gastroenterology 2019, 156 (7), 2024–2040. https://doi.org/10.1053/j.gastro.2019.01.259.

(11) Shi, C.; Fukushima, N.; Abe, T.; Bian, Y.; Hua, L.; Wendelburg, B. J.; Yeo, C. J.; Hruban, R. H.; Goggins, M. G.; Eshleman, J. R. Sensitive and Quantitative Detection of KRAS2 Gene Mutations in Pancreatic Duct Juice Differentiates Patients with Pancreatic Cancer from Chronic Pancreatitis, Potential for Early Detection. Cancer Biol. Ther. 2008, 7 (3), 353–360. https://doi.org/10.4161/cbt.7.3.5362.

(12) Long, E. E.; Van Dam, J.; Weinstein, S.; Jeffrey, B.; Desser, T.; Norton, J. A. Computed Tomography, Endoscopic, Laparoscopic, and Intra-Operative Sonography for Assessing Resectability of Pancreatic Cancer. Surg. Oncol. 2005, 14 (2), 105–113. https://doi.org/10.1016/j.suronc.2005.07.001.

(13) Sahani, D. V.; Shah, Z. K.; Catalano, O. A.; Boland, G. W.; Brugge, W. R. Radiology of Pancreatic Adenocarcinoma: Current Status of Imaging. J. Gastroenterol. Hepatol. 2008, 23 (1), 23–33. https://doi.org/10.1111/j.1440-1746.2007.05117.x.

(14) Vasen, H.; Ibrahim, I.; Robbers, K.; Van Mil, A. M.; Potjer, T.; Bonsing, B. A.; Bergman, W.; Wasser, M.; Morreau, H.; De Vos Tot Nederveen Cappel, W. H.; et al. Benefit of Surveillance for Pancreatic Cancer in High-Risk Individuals: Outcome of Long-Term Prospective Follow-up Studies from Three European Expert Centers. J. Clin. Oncol. 2016, 34 (17), 2010–2019. https://doi.org/10.1200/JCO.2015.64.0730.

(15) Jemal, A.; Center, M. M.; DeSantis, C.; Ward, E. M. Global Patterns of Cancer Incidence and Mortality Rates and Trends. Cancer Epidemiol. Biomarkers Prev. 2010, 19 (8), 1893–1907. https://doi.org/10.1158/1055-9965.EPI-10-0437.

(16) Lemaine, V.; Simmons, P. S. The Adolescent Female: Breast and Reproductive Embryology and Anatomy. Clin. Anat. 2013, 26 (1), 22–28. https://doi.org/10.1002/ca.22167.

(17) Pandya, S.; Moore, R. G. Breast Development and Anatomy. Clin. Obstet. Gynecol. 2011, 54 (1), 91–95. https://doi.org/10.1097/GRF.0b013e318207ffe9.

(18) Curtis, C.; Shah, S. P.; Chin, S. F.; Turashvili, G.; Rueda, O. M.; Dunning, M. J.; Speed, D.; Lynch, A. G.; Samarajiwa, S.; Yuan, Y.; et al. The Genomic and Transcriptomic Architecture of 2,000 Breast Tumours Reveals Novel Subgroups. Nature 2012, 486 (7403), 346–352. https://doi.org/10.1038/nature10983.

(19) Malhotra, G. K.; Zhao, X.; Band, H.; Band, V. Histological, Molecular and Functional Subtypes of Breast Cancers. Cancer Biol. Ther. 2010, 10 (10), 955–960. https://doi.org/10.4161/cbt.10.10.13879.

(20) Paap, E.; Verbeek, A. L. M.; Botterweck, A. A. M.; van Doorne-Nagtegaal, H. J.; Imhof-Tas, M.; de Koning, H. J.; Otto, S. J.; de Munck, L.; van der Steen, A.; Holland, R.; et al. Breast Cancer Screening Halves the Risk of Breast Cancer Death: A Case-Referent Study. Breast 2014, 23 (4), 439–444. https://doi.org/10.1016/j.breast.2014.03.002.

(21) Sankatsing, V. D. V. D. V.; Geuzinge, H. A. A.; Fracheboud, J.; Van Ravesteyn, N. T. T.; Heijnsdijk, E. A. M. A. M.; Kregting, L. M. M.; Broeders, M. J. M. J. M.; Otten, J. D. M. D. M.; Verbeek, A. L. M. L. M.; Pijnappel, R. .; et al. Landelijke Evaluatie van Bevolkingsonderzoek Naar Borstkanker in Nederland 2004 - 2014; 2019.

(22) Myers, E. R.; Moorman, P.; Gierisch, J. M.; Havrilesky, L. J.; Grimm, L. J.; Ghate, S.; Davidson, B.; Mongtomery, R. C.; Crowley, M. J.; McCrory, D. C.; et al. Benefits and Harms of Breast Cancer Screening: A Systematic Review. JAMA - J. Am. Med. Assoc. 2015, 314 (15), 1615–1634. https://doi.org/10.1001/jama.2015.13183.

(23) Shiovitz, S.; Korde, L. A. Genetics of Breast Cancer: A Topic in Evolution. Ann. Oncol. 2015, 26 (7), 1291–1299. https://doi.org/10.1093/annonc/mdv022.
(24) Wellings, E.; Vassiliades, L.; Abdalla, R. Breast Cancer Screening for High-Risk Patients of Different Ages and Risk - Which Modality Is Most Effective? Cureus 2016, 8 (12). https://doi.org/10.7759/cureus.945.
(25) Smith, L. M.; Kelleher, N. L. Proteoform: A Single Term Describing Protein Complexity. Nat. Methods 2013, 10 (3), 186–187. https://doi.org/10.1038/nmeth.2369.
(26) Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; et al. How Many Human Proteoforms Are There? Nat. Chem. Biol. 2018, 14 (3), 206–214. https://doi.org/10.1038/nchembio.2576.
(27) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J. Protein Posttranslational Modifications: The Chemistry of Proteome Diversifications. Angew. Chem. Int. Ed. Engl. 2005, 44 (45), 7342–7372. https://doi.org/10.1002/anie.200501023.
(28) Walsh, G. Post-Translational Modifications in the Context of Therapeutic Proteins: An Introductory Overview. In Post-translational Modification of Protein Biopharmaceuticals; Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2009; pp 1–14.
(29) Varki, A.; Esko, J. D.; Colley, K. J. Cellular Organization of Glycosylation. In Essentials of Glycobiology; Varki, A., Cummings, R., Esko, J., Freeze, H. H., Goldsmith, H. W., Bertozzi, C. R., Hart, G. W., E. Etzler, M., Eds.; Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2009.
(30) Varki, A. Biological Roles of Oligosaccharides: All of the Theories Are Correct. Glycobiology 1993, 3 (2), 97–130. https://doi.org/10.1093/glycob/3.2.97.
(31) Wormald, M. R.; Dwek, R. A. Glycoproteins: Glycan Presentation and Protein-Fold Stability. Structure 1999, 7 (7), R155-60.
(32) Geyer, H.; Geyer, R. Strategies for Analysis of Glycoprotein Glycosylation. Biochim. Biophys. Acta - Proteins Proteomics 2006, 1764 (12), 1853–1869. https://doi.org/10.1016/j.bbapap.2006.10.007.
(33) Gabius, H.-J. The Sugar Code: Fundamentals of Glycosciences, 2nd ed.; John Wiley & Sons, 2011.
(34) Helenius, A.; Aebi, M. Roles of N-Linked Glycans in the Endoplasmic Reticulum. Annu. Rev. Biochem. 2004, 73, 1019–1049. https://doi.org/10.1146/annurev.biochem.73.011303.073752.
(35) Brockhausen, I.; Schachter, H.; Stanley, P. O-GalNAc Glycans. In Essentials of Glycobiology; Varki, A., Cummings, R., Esko, J., Freeze, H. H., Goldsmith, H. W., Bertozzi, C. R., Hart, G. W., E. Etzler, M., Eds.; Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press: New York, 2009.
(36) Mariño, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. A Systematic Approach to Protein Glycosylation Analysis: A Path through the Maze. Nat. Chem. Biol. 2010, 6 (10), 713–723. https://doi.org/10.1038/nchembio.437.
(37) Varki, A. Biological Roles of Glycans. Glycobiology 2017, 27 (1), 3–49. https://doi.org/10.1093/glycob/cww086.
(38) Adamczyk, B.; Tharmalingam, T.; Rudd, P. M. Glycans as Cancer Biomarkers. Biochim. Biophys. Acta - Gen. Subj. 2012, 1820 (9), 1347–1353. https://doi.org/10.1016/j.bbagen.2011.12.001.
(39) Stowell, S. R.; Ju, T.; Cummings, R. D. Protein Glycosylation in Cancer. Annu. Rev. Pathol. Mech. Dis. 2015, 10 (1), 473–510. https://doi.org/10.1146/annurev-pathol-012414-040438.
(40) De Freitas-Junior, J. C. M.; Morgado-Díaz, J. A. The Role of N-Glycans in Colorectal Cancer Progression: Potential Biomarkers and Therapeutic Applications. Oncotarget 2016, 7 (15), 19395–19413. https://doi.org/10.18632/oncotarget.6283.
(41) Munkley, J. The Glycosylation Landscape of Pancreatic Cancer (Review). Oncol. Lett. 2019, 17 (3), 2569–2575. https://doi.org/10.3892/ol.2019.9885.
(42) Kizuka, Y.; Taniguchi, N. Enzymes for N-Glycan Branching and Their Genetic and Nongenetic Regulation in Cancer. Biomolecules 2016, 6 (2), 1–21. https://doi.org/10.3390/biom6020025.
(43) Dube, D. H.; Bertozzi, C. R. Glycans in Cancer and Inflammation - Potential for Therapeutics and Diagnostics. Nat. Rev. Drug Discov. 2005, 4 (6), 477–488. https://doi.org/10.1038/nrd1751.
(44) Arnold, J. N.; Saldova, R.; Abd Hamid, U. M.; Rudd, P. M. Evaluation of the Serum N-Linked Glycome for the Diagnosis of Cancer and Chronic Inflammation. Proteomics 2008, 8 (16), 3284–3293. https://doi.org/10.1002/pmic.200800163.
(45) Zhao, Y.; Sato, Y.; Isaji, T.; Fukuda, T.; Matsumoto, A.; Miyoshi, E.; Gu, J.; Taniguchi, N. Branched N-Glycans Regulate the Biological Functions of Integrins and Cadherins. FEBS J. 2008, 275 (9), 1939–1948. https://doi.org/10.1111/j.1742-4658.2008.06346.x.
(46) Zhang, Z.; Wuhrer, M.; Holst, S. Serum Sialylation Changes in Cancer. Glycoconj. J. 2018, 35 (2), 139–160. https://doi.org/10.1007/s10719-018-9820-0.
(47) Lan, Y.; Hao, C.; Zeng, X.; He, Y.; Zeng, P.; Guo, Z.; Zhang, L. Serum Glycoprotein-Derived N- and O-Linked Glycans as Cancer Biomarkers. Am. J. Cancer Res. 2016, 6 (11), 2390–2415.
(48) Reiding, K. R.; Vreeker, G. C. M.; Bondt, A.; Bladergroen, M. R.; Hazes, J. M. W.; van der Burgt, Y. E. M.; Wuhrer, M.; Dolhain, R. J. E. M. Serum Protein N-Glycosylation Changes with Rheumatoid Arthritis Disease Activity during and after Pregnancy. Front. Med. 2018, 4 (241), 1–11. https://doi.org/10.3389/fmed.2017.00241.
(49) Ruhaak, L. R.; Miyamoto, S.; Kelly, K.; Lebrilla, C. B. N-Glycan Profiling of Dried Blood Spots. Anal Chem 2012, 84 (1), 396–402. https://doi.org/10.1021/ac202775t.

(50) Holst, S.; Heijs, B.; De Haan, N.; Van Zeijl, R. J. M.; Briaire-De Bruijn, I. H.; Van Pelt, G. W.; Mehta, A. S.; Angel, P. M.; Mesker, W. E.; Tollenaar, R. A.; et al. Linkage-Specific in Situ Sialic Acid Derivatization for N-Glycan Mass Spectrometry Imaging of Formalin-Fixed Paraffin-Embedded Tissues. Anal. Chem. 2016, 88 (11), 5904–5913. https://doi.org/10.1021/acs.analchem.6b00819.

(51) Plomp, R.; De Haan, N.; Bondt, A.; Murli, J.; Dotz, V.; Wuhrer, M. Comparative Glycomics of Immunoglobulin A and G from Saliva and Plasma Reveals Biomarker Potential. Front. Immunol. 2018, 9 (OCT), 1–12. https://doi.org/10.3389/fimmu.2018.02436.

(52) Kammeijer, G. S. M.; Nouta, J.; De La Rosette, J. J. M. C. H.; De Reijke, T. M.; Wuhrer, M. An In-Depth Glycosylation Assay for Urinary Prostate-Specific Antigen. Anal. Chem. 2018, 90 (7), 4414–4421. https://doi.org/10.1021/acs.analchem.7b04281.

(53) Clerc, F.; Novokmet, M.; Dotz, V.; Reiding, K. R.; de Haan, N.; Kammeijer, G. S. M.; Dalebout, H.; Bladergroen, M. R.; Vukovic, F.; Rapp, E.; et al. Plasma N-Glycan Signatures Are Associated With Features of Inflammatory Bowel Diseases. Gastroenterology 2018, 155 (3), 829–843. https://doi.org/10.1053/j.gastro.2018.05.030.

(54) Battistel, M. D.; Azurmendi, H. F.; Yu, B.; Freedberg, D. I. NMR of Glycans: Shedding New Light on Old Problems. Prog. Nucl. Magn. Reson. Spectrosc. 2014, 79, 48–68. https://doi.org/10.1016/j.pnmrs.2014.01.001.

(55) Leeflang, B. R.; Faber, E. J.; Erbel, P.; Vliegenthart, J. F. G. Structure Elucidation of Glycoprotein Glycans and of Polysaccharides by NMR Spectroscopy. J. Biotechnol. 2000, 77 (1), 115–122. https://doi.org/10.1016/S0168-1656(99)00212-6.

(56) Lundborg, M.; Widmalm, G. Structural Analysis of Glycans by NMR Chemical Shift Prediction. Anal. Chem. 2011, 83 (5), 1514–1517. https://doi.org/10.1021/ac1032534.

(57) Taylor, M.E. & Drickamer, K. Introduction to Glycobiology, Third edit.; Oxford University Press: Oxford, 2006; Vol. 2nd ed.

(58) Wuhrer, M.; De Boer, A. R.; Deelder, A. M. Structural Glycomics Using Hydrophilic Interaction Chromatography (HILIC) with Mass Spectrometry. Mass Spectrom. Rev. 2009, 28 (2), 192–206. https://doi.org/10.1002/mas.20195.

(59) Stavenhagen, K.; Kolarich, D.; Wuhrer, M. Clinical Glycomics Employing Graphitized Carbon Liquid Chromatography–Mass Spectrometry. Chromatographia 2014, 78 (5–6), 307–320. https://doi.org/10.1007/s10337-014-2813-7.

(60) Ruhaak, L. R.; Deelder, A. M.; Wuhrer, M. Oligosaccharide Analysis by Graphitized Carbon Liquid Chromatography-Mass Spectrometry. Anal. Bioanal. Chem. 2009, 394 (1), 163–174. https://doi.org/10.1007/s00216-009-2664-5.

(61) Zauner, G.; Deelder, A. M.; Wuhrer, M. Recent Advances in Hydrophilic Interaction Liquid Chromatography (HILIC) for Structural Glycomics. Electrophoresis 2011, 32 (24), 3456–3466. https://doi.org/10.1002/elps.201100247.

(62) Vreeker, G. C. M.; Wuhrer, M. Reversed-Phase Separation Methods for Glycan Analysis. Anal. Bioanal. Chem. 2017, 409 (2), 359–378. https://doi.org/10.1007/s00216-016-0073-0.

(63) Melmer, M.; Stangler, T.; Premstaller, A.; Lindner, W. Comparison of Hydrophilic-Interaction, Reversed-Phase and Porous Graphitic Carbon Chromatography for Glycan Analysis. J. Chromatogr. A 2011, 1218 (1), 118–123. https://doi.org/10.1016/j.chroma.2010.10.122.

(64) Szabo, Z.; Guttman, A.; Rejtar, T.; Karger, B. L. Improved Sample Preparation Method for Glycan Analysis of Glycoproteins by CE-LIF and CE-MS. Electrophoresis 2010, 31 (8), 1389–1395. https://doi.org/10.1002/elps.201000037.

(65) Schwedler, C.; Kaup, M.; Weiz, S.; Hoppe, M.; Braicu, E. I.; Sehouli, J.; Hoppe, B.; Tauber, R.; Berger, M.; Blanchard, V. Identification of 34 N-Glycan Isomers in Human Serum by Capillary Electrophoresis Coupled with Laser-Induced Fluorescence Allows Improving Glycan Biomarker Discovery. Anal. Bioanal. Chem. 2014, 406 (28), 7185–7193. https://doi.org/10.1007/s00216-014-8168-y.

(66) Weiz, S.; Wieczorek, M.; Schwedler, C.; Kaup, M.; Braicu, E. I.; Sehouli, J.; Tauber, R.; Blanchard, V. Acute-Phase Glycoprotein N-Glycome of Ovarian Cancer Patients Analyzed by CE-LIF. Electrophoresis 2016, 37 (11), 1461–1467. https://doi.org/10.1002/elps.201500518.

(67) Gennaro, L. A.; Salas-Solano, O. On-Line CE-LIF-MS Technology for the Direct Characterization of N-Linked Glycans from Therapeutic Antibodies. Anal. Chem. 2008, 80 (10), 3838–3845. https://doi.org/10.1021/ac800152h.

(68) Liu, Y.; Salas-Solano, O.; Gennaro, L. A. Investigation of Sample Preparation Artifacts Formed during the Enzymatic Release of N-Linked Glycans Prior to Analysis by Capillary Electrophoresis. Anal. Chem. 2009, 81 (16), 6823–6829. https://doi.org/10.1021/ac9010588.

(69) Bunz, S. C.; Cutillo, F.; Neusüß, C. Analysis of Native and APTS-Labeled N-Glycans by Capillary Electrophoresis/Time-of-Flight Mass Spectrometry. Anal. Bioanal. Chem. 2013, 405 (25), 8277–8284. https://doi.org/10.1007/s00216-013-7231-4.

(70) Nakano, M.; Higo, D.; Arai, E.; Nakagawa, T.; Kakehi, K.; Taniguchi, N.; Kondo, A. Capillary Electrophoresis-Electrospray Ionization Mass Spectrometry for Rapid and Sensitive N-Glycan Analysis of Glycoproteins as 9-Fluorenylmethyl Derivatives. Glycobiology 2009, 19 (2), 135–143. https://doi.org/10.1093/glycob/cwn115.

(71) Mechref, Y.; Novotny, M. V. Glycomic Analysis by Capillary Electrophoresis-Mass Spectrometry. Mass Spectrom. Rev. 2009, 28 (2), 207–222. https://doi.org/10.1002/mas.20196.

(72) Lageveen-Kammeijer, G. S. M.; de Haan, N.; Mohaupt, P.; Wagt, S.; Filius, M.; Nouta, J.; Falck, D.; Wuhrer, M.

Highly Sensitive CE-ESI-MS Analysis of N-Glycans from Complex Biological Samples. Nat. Commun. 2019, 10 (1), 2137. https://doi.org/10.1038/s41467-019-09910-7.
(73) Lista, S.; Faltraco, F.; Hampel, H. Biological and Methodical Challenges of Blood-Based Proteomics in the Field of Neurological Research. Prog. Neurobiol. 2013, 101–102 (1), 18–34. https://doi.org/10.1016/j.pneurobio.2012.06.006.
(74) Dědová, T.; Grunow, D.; Kappert, K.; Flach, D.; Tauber, R.; Blanchard, V. The Effect of Blood Sampling and Preanalytical Processing on Human N-Glycome. PLoS One 2018, 13 (7), e0200507. https://doi.org/10.1371/journal.pone.0200507.
(75) Chambers, A. G.; Percy, A. J.; Yang, J.; Borchers, C. H. Multiple Reaction Monitoring Enables Precise Quantification of 97 Proteins in Dried Blood Spots. Mol. Cell. Proteomics 2015, 14 (11), 3094–3104. https://doi.org/10.1074/mcp.O115.049957.
(76) Merry, T.; Astrautsova, S. Chemical and Enzymatic Release of Glycans from Glycoproteins. Methods Mol. Biol. 2003, 213, 27–40. https://doi.org/10.1385/1-59259-294-5:27.
(77) Mulloy, B.; Dell, A.; Stanley, P.; Prestegard, J. H. Structural Analysis of Glycans. In Essentials of Glycobiology; Varki, A., Ed.; Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press: New York, 2009; pp 639–652.
(78) Rudd, P. M.; Karlsson, N. G.; Khoo, K.-H.; Packer, N. H. Glycomics and Glycoproteomics. In Essentials of Glycobiology; Varki, A., Ed.; Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press: New York, 2009; pp 653–666.
(79) Stanley, P.; Schachter, H.; Taniguchi, N. N-Glycans. In Essentials of Glycobiology; Varki, A., Cummings, R., Esko, J., Freeze, H. H., Goldsmith, H. W., Bertozzi, C. R., Hart, G. W., E. Etzler, M., Eds.; Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press: New York, 2009.
(80) Keser, T.; Pavić, T.; Lauc, G.; Gornik, O. Comparison of 2-Aminobenzamide, Procainamide and RapiFluor-MS as Derivatizing Agents for High-Throughput HILIC-UPLC-FLR-MS N-Glycan Analysis. Front. Chem. 2018, 6 (July), 1–12. https://doi.org/10.3389/fchem.2018.00324.
(81) Melmer, M.; Stangler, T.; Schiefermeier, M.; Brunner, W.; Toll, H.; Rupprechter, A.; Lindner, W.; Premstaller, A. HILIC Analysis of Fluorescence-Labeled N-Glycans from Recombinant Biopharmaceuticals. Anal. Bioanal. Chem. 2010, 398 (2), 905–914. https://doi.org/10.1007/s00216-010-3988-x.
(82) Ruhaak, L. R.; Steenvoorden, E.; Koeleman, C. A. M.; Deelder, A. M.; Wuhrer, M. 2-Picoline-Borane: A Non-Toxic Reducing Agent for Oligosaccharide Labeling by Reductive Amination. Proteomics 2010, 10 (12), 2330–2336. https://doi.org/10.1002/pmic.200900804.
(83) Lattova, E.; Perreault, H. Labelling Saccharides with Phenylhydrazine for Electrospray and Matrix-Assisted Laser Desorption-Ionization Mass Spectrometry. J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci. 2003, 793 (1), 167–179.
(84) Afiuni-Zadeh, S.; Rogers, J. C.; Snovida, S. I.; Bomgarden, R. D.; Griffin, T. J. AminoxyTMT: A Novel Multi-Functional Reagent for Characterization of Protein Carbonylation. Biotechniques 2016, 60 (4), 186–196. https://doi.org/10.2144/000114402.
(85) Gong, B.; Hoyt, E.; Lynaugh, H.; Burnina, I.; Moore, R.; Thompson, A.; Li, H. N-Glycosylamine-Mediated Isotope Labeling for Mass Spectrometry-Based Quantitative Analysis of N-Linked Glycans. Anal. Bioanal. Chem. 2013, 405 (17), 5825–5831. https://doi.org/10.1007/s00216-013-6988-9.
(86) Gennaro, L. A.; Harvey, D. J.; Vouros, P. Reversed-Phase Ion-Pairing Liquid Chromatography/Ion Trap Mass Spectrometry for the Analysis of Negatively Charged, Derivatized Glycans. Rapid Commun. Mass Spectrom. 2003, 17 (14), 1528–1534. https://doi.org/10.1002/rcm.1079.
(87) Kim, K.-J.; Kim, Y.-W.; Kim, Y.-G.; Park, H.-M.; Jin, J. M.; Hwan Kim, Y.; Yang, Y.-H.; Kyu Lee, J.; Chung, J.; Lee, S.-G.; et al. Stable Isotopic Labeling-Based Quantitative Targeted Glycomics (i-QTaG). Biotechnol. Prog. 2015, 31 (3), 840–848. https://doi.org/10.1002/btpr.2078.
(88) Walker, S. H.; Taylor, A. D.; Muddiman, D. C. Individuality Normalization When Labeling with Isotopic Glycan Hydrazide Tags (INLIGHT): A Novel Glycan-Relative Quantification Strategy. J. Am. Soc. Mass Spectrom. 2013, 24 (9), 1376–1384. https://doi.org/10.1007/s13361-013-0681-2.
(89) Kim, K.-J.; Kim, Y.-W.; Hwang, C.-H.; Park, H.-G.; Yang, Y.-H.; Koo, M.; Kim, Y.-G. A MALDI-MS-Based Quantitative Targeted Glycomics (MALDI-QTaG) for Total N-Glycan Analysis. Biotechnol. Lett. 2015. https://doi.org/10.1007/s10529-015-1881-6.
(90) Sekiya, S.; Wada, Y.; Tanaka, K. Derivatization for Stabilizing Sialic Acids in MALDI-MS. Anal. Chem. 2005, 77 (15), 4962–4968. https://doi.org/10.1021/ac050287o.
(91) Reiding, K. R.; Blank, D.; Kuijper, D. M.; Deelder, A. M.; Wuhrer, M. High-Throughput Profiling of Protein N-Glycosylation by MALDI-TOF-MS Employing Linkage-Specific Sialic Acid Esterification. Anal. Chem. 2014, 86 (12), 5784–5793. https://doi.org/10.1021/ac500335t.
(92) Harvey, D. J. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Carbohydrates. Mass Spectrom. Rev. 1999, 18 (6), 349–450. https://doi.org/10.1002/(SICI)1098-2787(1999)18:6<349::AID-MAS1>3.0.CO;2-H.
(93) Tsaprailis, G.; Nair, H.; Somogyi, Á.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. Influence of Secondary Structure on the Fragmentation of Protonated Peptides. J. Am. Chem. Soc. 1999, 121 (22), 5142–5154. https://doi.org/10.1021/ja982980h.
(94) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. a. Mobile and Localized Protons: A Framework for

Understanding Peptide Dissociation. J. Mass Spectrom. 2000, 35 (12), 1399–1406. https://doi.org/10.1002/1096-9888(200012)35:12<1399::AID-JMS86>3.0.CO;2-R.

(95) Wheeler, S. F.; Domann, P.; Harvey, D. J. Derivatization of Sialic Acids for Stabilization in Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry and Concomitant Differentiation of α (2 ⊚ 3)- and α (2 ⊚ 6)-Isomers. Rapid Commun. Mass Spectrom. 2009, 23 (2), 303–312. https://doi.org/10.1002/rcm.3867.

(96) Bladergroen, M. R.; Reiding, K. R.; Hipgrave Ederveen, A. L.; Vreeker, G. C. M.; Clerc, F.; Holst, S.; Bondt, A.; Wuhrer, M.; van der Burgt, Y. E. M. Automation of High-Throughput Mass Spectrometry-Based Plasma N -Glycome Analysis with Linkage-Specific Sialic Acid Esterification. J. Proteome Res. 2015, 14 (9), 4080–4086. https://doi.org/10.1021/acs.jproteome.5b00538.

(97) Alley, W. R.; Novotny, M. V. Glycomic Analysis of Sialic Acid Linkages in Glycans Derived from Blood Serum Glycoproteins. J. Proteome Res. 2010, 9 (6), 3062–3072. https://doi.org/10.1021/pr901210r.

(98) Toyoda, M.; Ito, H.; Matsuno, Y. K.; Narimatsu, H.; Kameyama, A. Quantitative Derivatization of Sialic Acids for the Detection of Sialoglycans by MALDI MS. Anal. Chem. 2008, 80 (13), 5211–5218. https://doi.org/10.1021/ac800457a.

(99) Alley, W. R.; Madera, M.; Mechref, Y.; Novotny, M. V. Chip-Based Reversed-Phase Liquid Chromatography-Mass Spectrometry of Permethylated N-Linked Glycans: A Potential Methodology for Cancer-Biomarker Discovery. Anal. Chem. 2010, 82 (12), 5095–5106. https://doi.org/10.1021/ac100131e.

(100) Delaney, J.; Vouros, P. Liquid Chromatography Ion Trap Mass Spectrometric Analysis of Oligosaccharides Using Permethylated Derivatives. Rapid Commun. Mass Spectrom. 2001, 15 (5), 325–334. https://doi.org/10.1002/rcm.230.

(101) Ng, E. W. Y.; Wong, M. Y. M.; Poon, T. C. W. Advances in MALDI Mass Spectrometry in Clinical Diagnostic Applications. In Chemical Diagnostics From Bench to Bedside; Tang, N. L. S., Poon, T., Eds.; Springer, 2013; pp 139–175. https://doi.org/10.1007/128_2012_413.

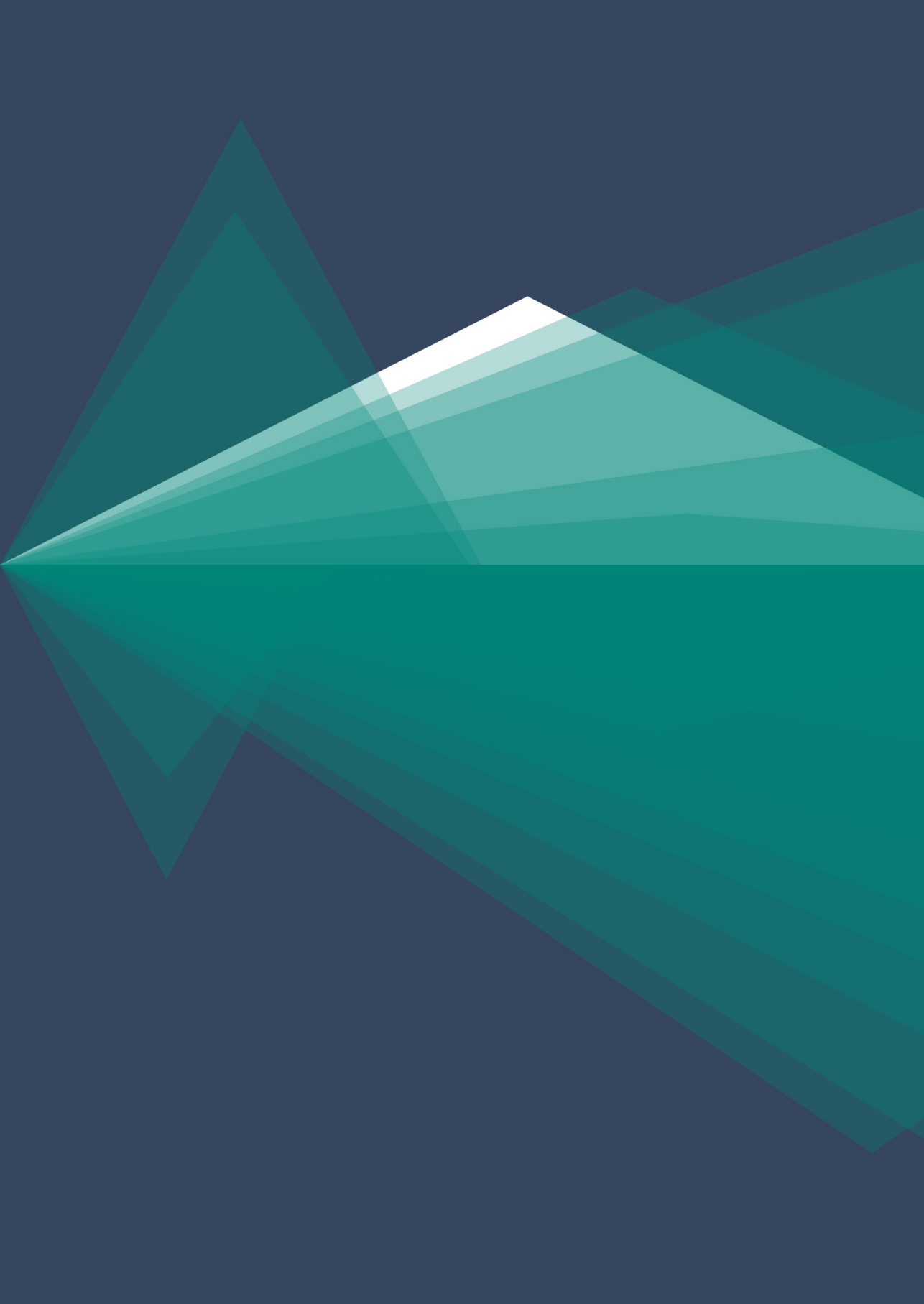
(102) Reiding, K. R.; Lonardi, E.; Hipgrave Ederveen, A. L.; Wuhrer, M. Ethyl Esterification for MALDI-MS Analysis of Protein Glycosylation. Methods Mol. Biol. 2016, 1394, 151–162. https://doi.org/10.1007/978-1-4939-3341-9_11.

(103) de Hoffmann, E.; Stroobant, V. Mass Spectrometry: Principles and Applications, Third.; John Wiley & Sons, 2007.

(104) Zenobi, R.; Zenobi, R.; Knochenmuss, R.; Knochenmuss, R. Ion Formation in MALDI Mass Spectrometry. Mass Spectrom. Rev. 1998, 17 (5), 337–366. https://doi.org/10.1002/(SICI)1098-2787(1998)17:5<337::AID-MAS2>3.0.CO;2-S.

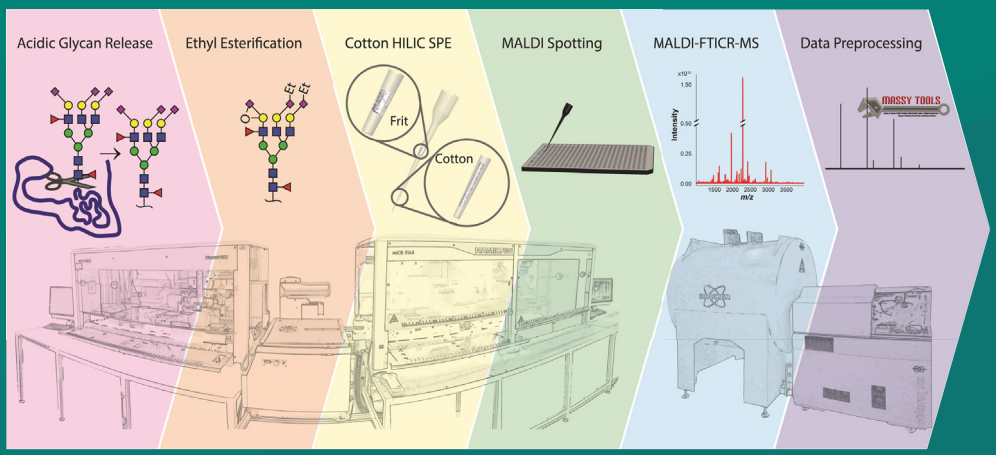
(105) Albrethsen, J. Reproducibility in Protein Profiling by MALDI-TOF Mass Spectrometry. Clin. Chem. 2007, 53 (5), 852–858. https://doi.org/10.1373/clinchem.2006.082644.

(106) Krutchinsky, A. N.; Chait, B. T. On the Nature of the Chemical Noise in MALDI Mass Spectra. J. Am. Soc. Mass Spectrom. 2002, 13 (2), 129–134. https://doi.org/10.1016/S1044-0305(01)00336-1.



CHAPTER 2

Automated plasma glycomics with linkage-specific sialic acid esterification and ultrahigh resolution MS



Reprinted with permission from:

Vreeker, G.C.M.[1,2], *Nicolardi, S.*[1,2], *Bladergroen, M.R.*[1], *van der Plas, C.J.*[1], *Mesker, W.E.*[2], *Tollenaar, R.A.E.M.*[2], *van der Burgt, Y.E.M.*[1], *Wuhrer, M.*[1]*: Automated Plasma Glycomics with Linkage-Specific Sialic Acid Esterification and Ultrahigh Resolution MS. Anal. Chem. 90, 11955–11961 (2018). doi:10.1021/acs.analchem.8b02391. Supplemental info available.*

1 Center for Proteomics and Metabolomics, Leiden University Medical Center, The Netherlands

2 Department of Surgery, Leiden University Medical Center, The Netherlands

Abstract

High throughput mass spectrometry (MS) glycomics is an emerging field driven by technological advancements including sample preparation and data processing. Previously, we reported an automated protocol for the analysis of *N*-glycans released from plasma proteins that included sialic acid derivatization with linkage-specificity, namely ethylation of α2,6-linked sialic acid residues and lactone formation of α2,3-linked sialic acids. In the current study, each step in this protocol was further optimized. Method improvements included minimizing the extent of side-reaction during derivatization, an adjusted glycan purification strategy and mass analysis of the released *N*-glycans by ultrahigh resolution matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance MS. The latter resolved peak overlap and simplified spectral alignment due to high mass measurement precision. Moreover, this resulted in more confident glycan assignments and improved signal-to-noise for low-abundant species. The performance of the protocol renders high-throughput applications feasible in the exciting field of clinical glycomics.

Introduction

Without compromising the value of genomic and transcriptomic analyses it is widely acknowledged that the complexity of the human body largely arises from variations in protein expression as well as modifications between cells, tissues and body fluids. In this context, mass spectrometry (MS)-based proteomics has greatly contributed to an understanding of cellular functions at a molecular level.[1] Interestingly, in MS-based biomarker discovery efforts protein glycosylation has often not been taken into account, albeit that glycoproteins are key players in biological processes such as cell adhesion, endocytosis, receptor activation, signal transduction, molecular trafficking, and clearance.[2,3] Moreover, glycosylation has proven to be reflective of (and in some cases causal to) disease etiology. Notable examples include tumor growth and metastasis, various inflammatory conditions and autoimmune diseases, and congenital disorders of glycosylation.[4,5] This neglect of glycans, and proteoforms in general, can be rationalized from data complexity and technical challenges.[6–8] In MS-based proteomics, database searches methyl-, phosphate- and acetyl-groups are accommodated by an exactly defined mass difference, whereas glycans have different sizes (monomeric to oligomeric) that moreover can be isobaric. With regard to technical considerations, the applied analytical methods require high-throughput (HT) platforms to study the clinical relevance of protein glycosylation in large-scale studies, as larger sample sizes have a higher risk of errors and inconsistencies in sample preparation, leading to poor repeatability.[9]

Multiple HT methods have been developed for proteomics[10,11] and glycomics[9,12–21] purposes. For released *N*-glycan analysis several HT methods are available: capillary electrophoresis with laser-induced fluorescence detection (CE-LIF)[14,15], ultra-performance liquid chromatography with fluorescence detection (UPLC-FLD)[16,18], and matrix-assisted laser desorption/ionization mass spectrometry (MALDI)-MS.[12,22] Some of these methods include LC runs which show a relatively lower throughput compared to MALDI-MS and multiplexed CE-LIF. In addition, fluorescence detection is often used, which has limitations for highly complex samples due to potentially overlapping signals. Here we present a high-performance mass spectrometric glycomics assay which includes automated sample preparation and high resolution mass spectrometry. This assay shows a range of advantages as compared to other workflows and is based on our previous plasma glycomics protocol that includes sialic acid derivatization with linkage-specificity and a plate-based method for glycan purification. High throughput: The assay can be run at high throughput allowing the measurement of 576 samples within one day, thereby making the analysis of large-scale clinical cohorts possible. High coverage: 120 glycans were detected from human plasma applying linkage-specific sialic acid

derivatization, 90 glycans whereof were consistently quantified. High confidence identification: artifacts and contaminants were minimized by optimizing the sample preparation. Together with the high-resolution, high mass accuracy and precision, this tackled issues regarding signal overlap and allowed confident glycan identification. High performance relative quantification: Due to the automated and optimized sample preparation including cotton thread micro-solid phase extraction tips, reproducible glycan profiles were obtained. This novel high-performance workflow allows the robust performance of large-scale clinical glycomics studies.

Experimental section

Samples

Plasma standard (Visucon-F frozen normal control plasma, pooled from twenty human donors, citrated and buffered with 0.02 M HEPES) was purchased from Affinity Biologicals (Ancaster, ON, Canada) and used for all experiments.

Chemicals, reagents and enzymes

Analytical grade ethanol, sodium dodecyl sulfate (SDS), trifluoroacetic acid (TFA) and potassium hydroxide (KOH) were obtained from Merck (Darmstadt, Germany). Disodium hydrogen phosphate dihydrate ($Na_2HPO_4 \times 2\,H_2O$), potassium dihydrogen phosphate (KH_2PO_4), sodium chloride (NaCl), N,N′-diisopropylcarbodiimide (DIC), N,N′-dicyclohexylcarbodiimide (DCC), 85% phosphoric acid (H_3PO_4), 50% sodium hydroxide (NaOH), nonidet P-40 substitute (NP-40), 1-hydroxybenzotriazole 97% (HOBt) and super-DHB (9:1 mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid, sDHB) were purchased from Sigma-Aldrich (Steinheim, Germany). 1-ethyl-3-(3-(dimethylamino)propyl)carbodiimide (EDC) hydrochloride was obtained from Fluorochem (Hadfield, UK), whereas recombinant peptide-*N*-glycosidase F (PNGase F) was obtained from Roche Diagnostics (Mannheim, Germany) and HPLC-grade acetonitrile (ACN) was purchased from Biosolve (Valkenswaard, The Netherlands). Milli-Q water (MQ) was generated from a Q-Gard 2 system (Millipore, Amsterdam, the Netherlands), which was maintained at ≥18 MΩ.

Enzymatic *N*-glycan release

The enzymatic release of *N*-glycans from plasma proteins was performed as previously described, however using an acidified PBS buffer.[12,23–25] Briefly, acidified PBS (pH 5.6) was prepared by adding 68 μL of 85% phosphoric acid to 9.93 mL 5x PBS solution. Then, 6 μL of plasma was added to 12 μL of 2% SDS and incubated at 60 °C for 10 minutes. Next, 12.6 μL of freshly prepared releasing mixture (6 μL of 4% NP-40, 6 μL of acidified PBS and 0.6 μL of PNGase F) was added and the samples were overnight incubated at 37 °C. Thus prepared samples were stored at -20 °C until further analysis and measurement.

Automated sample preparation

The derivatization, hydrophilic interaction liquid chromatography (HILIC) purification and MALDI-target plate spotting were performed using an automated liquid handling platform as described in Bladergroen et al.[12] This platform allows for processing of 96 samples (i.e. one microtitration plate (MTP)) in a simultaneous manner

and can accommodate sequential processing of six MTPs without interruption. The first plate requires 1.5 hours for processing, whereas the following ones each add half an hour to the total processing time due to efficient (overlapping) time scheduling.[12] The ethyl esterification derivatization was performed by adding 2 μL of released glycan sample to 40 μL ethyl esterification reagent (0.25 M EDC with 0.25 HOBt in ethanol) and incubating the mixture for one hour at 37 °C.[24] Subsequently 40 μL of acetonitrile was added and after 10 minutes the purification was started. In-house assembled microtips used for cotton HILIC microtip purification were prepared as follows: 3 mm cotton thread (approximately 180 μg, Pipoos, Utrecht, Netherlands) was placed into a 50 μL tip (clear CO-RE tip without filter, Hamilton, Switzerland) by using tweezers. Then, a polypropylene porous frit (DPX Technologies, Columbia GA, United States of America) was placed 18 mm above the tip opening in order to prevent the cotton thread from floating through the microtips during purification. The cotton HILIC tips were three times pre-wetted with 40 μL of MQ water and then conditioned with three times 40 μL 85% ACN. Subsequently the sample was loaded by pipetting the ethyl-esterified sample twenty times up and down (40 μL per time). The HILIC tips were washed three times with 40 μL 85% ACN containing 1% TFA, and three times with 40 μL 85% ACN. The purified *N*-glycans were eluted in 20 μL MQ water by pipetting five times up and down (pipet set at 15 μL). Next, 10 μL of purified sample was pre-mixed with 5 μL of sDHB matrix (5 mg/mL in 99% ACN with 1 mM NaOH) and 3 μL of the mixture was spotted onto a MALDI target plate (800/384 MTP AnchorChip, Bruker Daltonics, Bremen, Germany). The spots were allowed to dry in air, followed by MALDI-time-of-flight (TOF)-MS measurement and subsequent MALDI-FTICR-MS measurement of the exact same spot.

Sample work-up for studying adduct formation

The *N*-glycan release was performed three times in acidified PBS (pH 5.6) and nine times in "regular" 5x PBS (pH 7.3). Ethyl esterification was carried out according to the protocol described above, however for the group of nine samples the reaction was performed three times with three different activating agents: 0.25 M EDC, 0.25 M DCC or 0.25 M DIC. The samples were subsequently purified with the manual cotton HILIC solid phase extraction (SPE) microtips as described previously[24,26], spotted onto a MALDI target plate and allowed to dry in air.

MALDI-TOF-MS

MALDI-TOF-MS spectra were recorded in reflectron positive mode on an UltrafleXtreme mass spectrometer (Bruker Daltonics). The system was equipped with a Smartbeam-II laser and operated by flexControl version 3.4 Build 135. Per sample 10,000 laser shots were collected at a laser frequency of 1000 Hz. The spot

was analyzed using a random walking pattern of 200 shots per raster spot. Spectra were recorded with an *m*/*z*-range of 1,000 to 5,000.

MALDI-FTICR-MS and MS/MS

MALDI-FTICR-MS measurements were performed on a Bruker 15T solariX XR FTICR mass spectrometer equipped with a CombiSource and a ParaCell (Bruker Daltonics). The system was controlled by ftmsControl version 2.1.0 and spectra were recorded with 1M data points (i.e., transient length of 2.307 s). A Bruker Smartbeam-II Laser System was used for irradiation at a frequency of 500 Hz using the "medium" predefined shot pattern. Each MALDI-FTICR spectrum was generated in the *m*/*z*-range from 1011.86 to 5000.00. At each raster 200 laser shots were collected and each spectrum consisted of ten acquired scans. The ParaCell parameters were as follows: the DC bias RX0, TX180, RX180, and TX0 were 9.13, 9.20, 9.27, 9.20 V, respectively; the trapping potentials were set at 9.50 and 9.45 V and the excitation power and sweep step time at 55% and 15 µs. The transfer time of the ICR cell was 1.0 ms, and the quadrupole mass filter was set at *m*/*z* 850. Collision-induced dissociation (CID) experiments were performed in the *m*/*z*-range 153.34-5000 with 512 k data points. The quadrupole (Q1) was used from precursor ion selection with an isolation window of 12 mass units. Fragments were generated using a collision energy of 75 V. DataAnalysis Software 4.2 (Bruker Daltonics) was used for the visualization and data analysis of MALDI-(CID)-FTICR spectra.

Method repeatability

The repeatability of the method was evaluated by performing the sample preparation and measurements multiple times on three different days. On day 1 the process described in the section 'automated sample preparation' was repeated three times with 96 samples each time and on day 2 and 3 the process was repeated once with 96 samples per day. From this data the intra-plate, intra-day and inter-day variation can be determined. The intra-plate variation was calculated over the total amount of spectra per plate. The intra-day variation was calculated over the total amount of spectra from day 1 (plate 1-3) and the inter-day variation was calculated over the total amount of spectra from all days (plate 1-5).

Preprocessing of mass spectral data

The MALDI-TOF-MS spectra were transformed in text format (x,y) using flexAnalysis 3.4 (Bruker Daltonics). The data transformation to text format of the MALDI-FTICR-MS spectra was performed with DataAnalysis 3.1 (Bruker Daltonics). With MassyTools 0.1.8.1, the text files were calibrated if at least five calibrants showed a signal-to-noise (S/N) ratio of 9 or higher, which was the case in all 480 MALDI-TOF-

MS spectra and 479 out of 480 MALDI-FTICR-MS spectra. The list of calibration masses is shown in the Supporting Information Table S-1. In addition, the spectra showing less than (average of all spectra) minus 3*(standard deviation) of the analyte area above S/N 9 were excluded in the analysis (13 MALDI-TOF-MS spectra and 6 MALDI-FTICR-MS spectra).

In the obtained spectra, the *m*/*z* values corresponding to 163 glycan compositions including 19 "dummy signals" (as defined in the supporting information Table S-2) were integrated with a coverage of 95% of the theoretical isotopic envelope per analyte. For the MALDI-TOF-MS data an extraction window of 0.49 *w*-units was applied, while for the MALDI-FTICR-MS spectra an increasing mass window was used for increasing *m*/*z* values (see supporting information Table S-3). Analyte curation was performed with multiple quality criteria: the S/N of the analyte for the median of all spectra had to be at least 6, the quality score for the comparison of the theoretical and observed isotopic envelope of the analyte had to be lower than 0.5 for the median of all spectra and the PPM-error of the analyte had to be lower than 10 for the first quartile of all spectra. All dummy analytes were excluded by these criteria. In addition, 64 signals were excluded for further analysis in the MALDI-TOF-MS spectra and 53 in the MALDI-FTICR-MS spectra. The complete list of *N*-glycans included in the extraction is shown in Table S-4 of the Supporting Information. In this list also the observed glycan analytes that could not be extracted because of lacking intensity or overlap with other signals are displayed. For MALDI-TOF-MS and MALDI-FTICR-MS spectra these were respectively 5 and 19 analytes. Additionally, for 10 analytes only the first isotopic peak could be used for analysis, due to overlap of the other isotopic peaks with other signals.

The extracted area values for the selected glycans were normalized to the sum of all areas for each spectrum separately. In addition derived traits were calculated from single glycans using R-Studio software with an in-house written script. The exact calculations for these derived traits, including the information on which glycans are involved in a certain trait, are shown in supporting information Table S-5.

Per analyte and trait the average area, standard deviation (SD) and relative standard deviation (RSD) were calculated. In addition, the average RSD of the 25 most abundant glycans was determined. These calculations were done for the intra-plate, intra-day and inter-day variation.

Results and discussion

Protocol development

The initial plasma glycomics protocol that includes a sialic acid derivatization method with linkage-specificity was used as a starting point for the development and optimization of a high performance assay.[12] Method improvements included minimizing the extent of side-reaction during derivatization, an adjusted glycan purification strategy and mass analysis of the released *N*-glycans by MALDI-FTICR. The benefits of MALDI-FTICR mass analysis of the plasma glycome will be presented and compared to MALDI-TOF readout. With these improvements the assay is now suitable for HT studies with predefined quality metrics.[27,28] A resulting ultrahigh resolution glycan profile of the improved and optimized protocol is exemplified in Figure 1. Here, released *N*-glycans were accurately identified up to *m*/*z*-values higher than 4000. Moreover, high precision measurement of mass differences allowed identification of a byproduct as follows

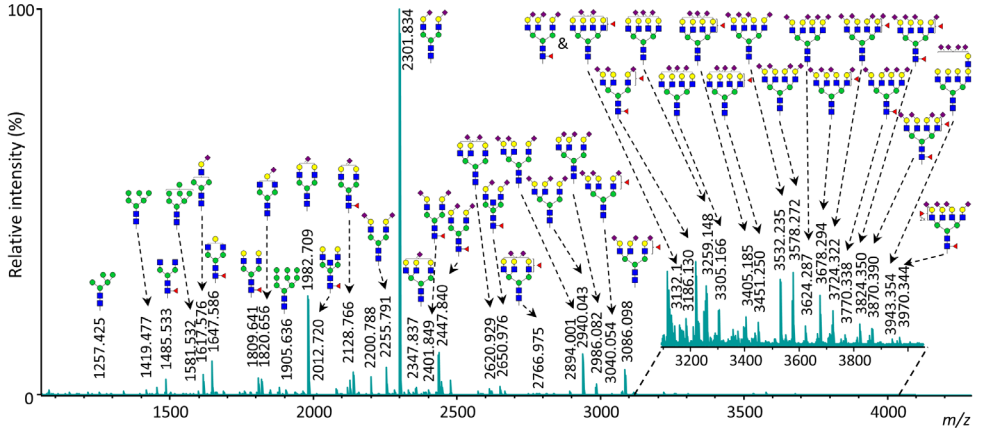


Figure 1. Annotated MALDI-FTICR-MS spectrum of released total plasma *N*-glycome, with the signals assigned as $[M+Na]^+$ compositions. The glycans were enzymatically released and subsequently ethyl esterified, cotton HILIC purified and spotted on a MALDI-target by the automated pipetting platform. The linkages and positions of the monosaccharides are based on literature as linkage specific information cannot be extracted from the MS-data, except for the *N*-acetylneuraminic acid monosaccharides of which the linkage is determined by the ethyl esterification derivatization method.

Identification of a byproduct from derivatization

Upon careful inspection of the MALDI-based glycan profiles (such as depicted in Figure 1), it was observed that satellites of annotated glycans appeared with mass increases of +132.174 ± 0.002 Da and 357.205 ± 0.029 Da, respectively (averaged over 114 profiles for three different analytes). Previously, these were explained as

unidentified reducing end modifications or ionization variants, namely at +132.19 Da and +357.18 Da of each glycan signal.[29] These signals only appear when the ethyl esterification method is used, so it is likely that one of the chemicals used in this reaction reacts with the glycan. In Figure 2A the hypothesized reaction of EDC and other carbodiimides with the reducing end of the glycan is illustrated. For this reaction it is assumed that the hydrolysis reaction of the glycosylamine after PNGase F release has not completed before addition of the ethyl esterification reagent. The resulting molecule matches the observed *m*/*z*-values when it is ionized as $[M+H]^+$. To further confirm this hypothesis, fragmentation of the +132.19-signal of H5N4E2 at *m*/*z* 2434.002 was performed and showed that the additional mass was indeed located at the reducing end and ionized as a proton adduct (Figure 2B and 2C).

The observed fragments (see also supporting information Figure S-1) match the structure as it was proposed in Figure 2A. This reaction results in a theoretical mass shift of 132.1763 *m*/*z*-units, which matches the observed mass shift. To proof that it is truly an addition of the complete EDC molecule, the reaction was also performed with two carbodiimides with different side groups. For these samples the glycan release was performed at regular pH 7.3. The results of these experiments are shown in Figure 2D panel 1-3. The shift of the extra peaks corresponds exactly to the mass differences between the different carbodiimides. In order to test whether the artifact is linked to the presence of glycosylamines, we performed the glycan release at slightly acidic conditions, knowing that glycosylamines are sensitive to acidic hydrolysis. The glycans from the sample in Figure 2D panel 4 were released with PNGase F in acidified PBS and ethyl esterified with EDC as an activating reagent. In this spectrum no signals corresponding to the byproduct were visible. This is explained by the hydrolysis reaction that is favored at lower pH. This adjustment of the pH of the glycan release reaction prevents the presence of glycosylamines in the subsequent ethyl esterification reaction. In addition, it is worth noting that the signals at +357.205 Da also disappear from the spectra when the glycan release is performed at lower pH. This indicated that these satellites may indeed be linked to the presence of glycosylamines. Interestingly, these peaks do not change when using different carbodiimides, implying that the additional mass originates from another source. Clearly, the "removal" of these satellites of annotated glycans by performing the release in acidified PBS simplifies the MALDI-profiles. It is furthermore stressed that the here applied FTICR-MS read-out facilitates the identification of other, obviously unwanted, satellite or unexplained signals. For example, in MALDI-based analysis of glycans the potential loss of the reducing end GlcNAc (-221.0899 Da or in the case of a core fucosylation -367.1478 Da) should be taken into account, as well as signals resulting from a 0,2A cross-ring fragmentation[29] or a 2,4A cross-ring fragmentation of the reducing end GlcNAc.[29]

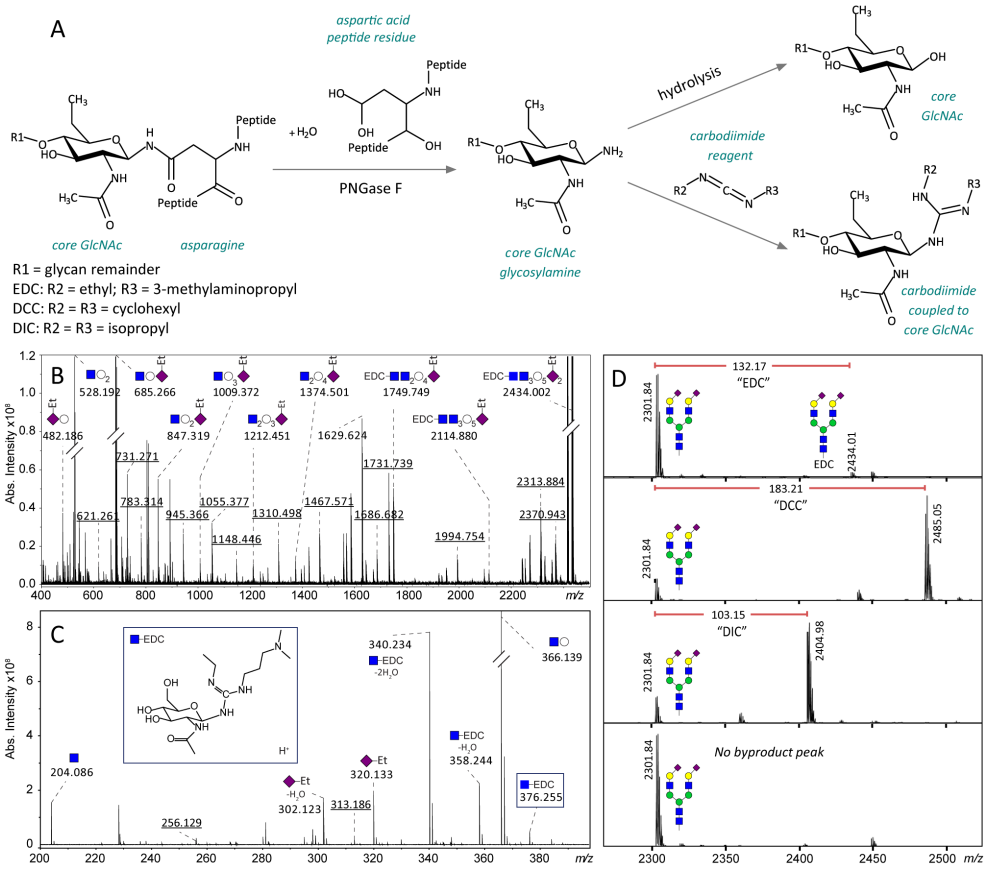


Figure 2. A) Reaction scheme of the observed reaction of the reducing end glycosylamine with a carbodiimide activating reagent. Chemical structures and reactions were drawn in Marvin (Marvin 17.26.0, ChemAxon). B and C) MALDI-CID-FTICR-MS spectrum of EDC-linked H5N4E2 selected at *m/z* 2434.002. For more fragmentation spectra, see supporting information Figure S-1. D) MALDI-TOF-MS spectra of derivatized H5N4E2. In panel D1-3 the observed *m/z* difference between the free reducing end glycan (H5N4E2 at *m/z* 2301.84) and the same glycan with the different carbodiimides attached to the reducing end is shown. The difference in peak shift corresponds to the mass difference of the R-groups of the different carbodiimides. For these samples, the glycan release was performed at regular pH 7.3. In panel D4 a spectrum is shown of a glycan release reaction performed with acidified PBS including ethyl esterification activated by EDC. In this spectrum, no signals corresponding to bound EDC are found.

Automated cotton HILIC purification

It was found that our initial GHP plate-based HILIC-SPE method for glycan purification (GHP plate, Pall AcroPrep Advance 96 Filter plate, Pall Corporation, Ann Arbor, MI) suffered from batch-to-batch differences for released glycans.[12] Both

the signal intensity in the spectra and the number of glycans visible in the spectra were decreasing over time. A clear explanation for this decrease could not be traced and for that reason an alternative purification method was added to the protocol. Here, automated HILIC purification was performed with cotton thread as the stationary phase that previously has been used in manual purification protocols.[24] For the automated method a piece of cotton thread was put into a pipet tip and was kept in place by capping it with a porous frit. These newly and in-house prepared microtips enabled the application of cotton HILIC tips in an automated pipetting platform.

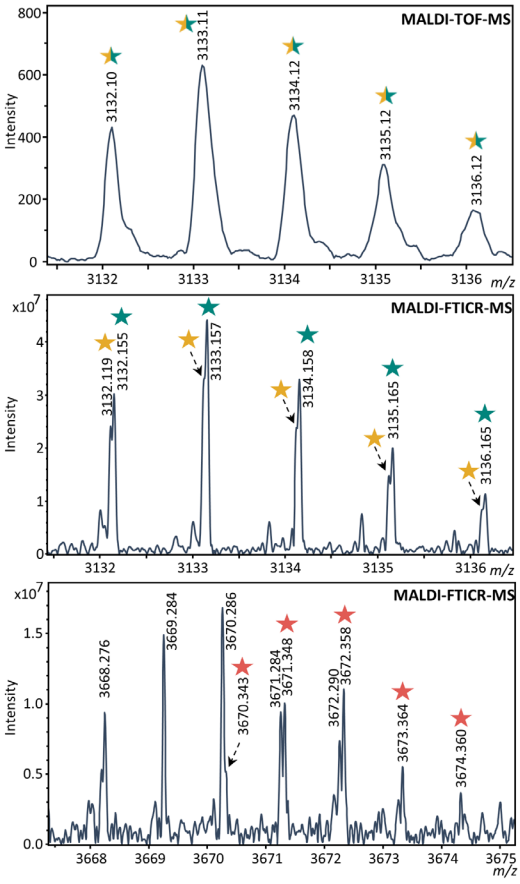
Using the automated cotton HILIC tips for purification instead of a GHP filter plate the spectral quality increased with respect to the signal intensity and glycan coverage. Moreover, using purification tips instead of a filter plate implies that there is no need to have a vacuum manifold or centrifuge (for elution) available on the pipetting platform. Furthermore, the overall throughput of the protocol is increased.

Glycan assignments

As the resolving power in MALDI-FTICR-MS spectra is higher compared to MALDI-TOF (ranging from 20,000 to 10,000 in TOF versus 100,000 to 50,000 in FTICR[30]), overlapping species in the glycan profiles become visible (see top and middle panel of Figure 3). The narrow peaks result in a higher peak capacity in the spectrum and moreover are measured at higher precision compared to MALDI-TOF-MS data. With suitable calibration of the MALDI-FTICR-MS spectra (<5 ppm errors for calibrants), the accuracy of the measurement increases to a level where the measured *m*/*z* can give a strong indication of the monosaccharide composition of the analyte. As a consequence, fewer false assignments are made and annotations are of higher confidence. In addition, the number of detected glycan signals is increased due to the larger dynamic range of the MALDI-FTICR-MS: signals with intensities near the limit of detection in MALDI-TOF-MS spectra could not be included for analysis, but the corresponding glycan signals in MALDI-FTICR-MS spectra showed higher intensity and S/N ratio, which made it possible to include these signals in the analysis. More details on the observed and relatively quantified analytes in MALDI-TOF-MS versus MALDI-FTICR-MS are found in supporting information Table S-4.

In the MALDI-TOF-MS spectra 86 glycan signals were annotated of which 80 could be extracted for relative quantification. For the MALDI-FTICR-MS spectra these numbers were 112 and 91, respectively. From these quantifiable glycans 76 were extracted from both MALDI-TOF-MS and MALDI-FTICR-MS spectra, 15 glycan signals were exclusively extracted in MALDI-FTICR-MS and 4 were exclusively extracted in MALDI-TOF-MS spectra. It should be noted that of these four MALDI-

TOF-MS glycan signals (see supporting information Table S-4) one (H7N6E4) was observed in MALDI-FTICR-MS spectra, but due to overlap could not be quantified, see Figure 3. The three remaining glycan signals did not have a corresponding signal in MALDI-FTICR-MS spectra, in fact it was observed that the isotopic pattern of these signals did not correspond to glycan identities.



Symbol	Composition	Theoretical m/z	Depiction
★	H7N6F1L1E1	3132.11526	
★	H6N5F1E3	3132.15165	
★	H7N6E4	3670.35265	

Figure 3. Resolving power of MALDI-FTICR-MS. In the top MALDI-TOF-MS spectrum a single isotopic pattern is visible, while in the middle MALDI-FTICR-MS spectrum there are two isotopic patterns observed. In the lower MALDI-FTICR-MS spectrum the glycan signal of H7N6E4 can be distinguished from an overlapping signal, but could not be extracted for relative quantification.

As a result of the higher resolving power of MALDI-FTICR-MS measurements the glycan signals were annotated with higher confidence. For most of the annotated glycans we published in previous studies the current data corroborated with proposed compositions. However, we also found a few interesting exceptions. In both MALDI-TOF-MS and MALDI-FTICR-MS a signal was observed around *m/z*-value 3132.14 (see Figure 3). In previous MALDI-TOF-MS studies this signal was often annotated as H6N5F1E3, a tri-antennary glycan which matches all quality criteria of the data preprocessing. However, in MALDI-FTICR-MS spectra not one, but two partially overlapping signals are observed, as is shown in Figure 3. One of these signals indeed matches the mass of the triantennary glycan composition stated above, but the other signal can very well be annotated as a signal corresponding to the tetraantennary glycan H7N6F1L1E1. In MALDI-TOF-MS it seemed that the signal was coming from one analyte, while the resolution of the MALDI-FTICR-MS spectra show that this signal is actually made up from two different glycan species. Notably there were also some cases in which the previous annotation was proven wrong. For example, a species previously assigned as H4N6F1E1 (*m/z*-value 2372.87) on the basis of TOF data[12] is now considered to be a byproduct or adduct from another glycan, because the signal highly decreases when the glycan release is performed at lower pH. In addition, there were some signals which in MALDI-TOF-MS were annotated with some compositions such as H4N7E1 that are questionable in a biosynthetic manner. Here the high resolution in combination with a good calibration showed that the ppm error for this composition and signal was much higher than for the other signals.

Repeatability of glycan relative quantifications

From the obtained MALDI-TOF-MS as well as MALDI-FTICR-MS spectra the variation of the sample preparation combined with the measurements was calculated for both systems. The intra-plate-, intra-day- and inter-day variation was calculated for the glycan with the highest relative area (H5N4E2) (see Supporting Information Table S-6) and averaged for the ten glycans (specific compositions are shown in Supporting Information Table S-6) with the highest relative area, which is shown in Figure 4. In this figure also the variation of the 25 glycans with the highest relative area is shown. The variation of all glycans is shown in supporting information Figure S-2 and in addition the variation of the calculated derived traits is shown in supporting information Figure S-3. From the information presented in Figure 4 and supporting information Table S-7 we can conclude that the variation of the MALDI-TOF-MS and MALDI-FTICR-MS measurements is similar to those reported previously: the variation of the most abundant signal is found to be around 5%.[12] For the ten most abundant signals the average intra-plate variation is 14.9% for the MALDI-TOF-MS

measurements, whereas for the MALDI-FTICR-MS measurements this variation is only 9.0%. When comparing the inter-day variation this slight difference is also visible with 16.5% and 9.9% variation for the MALDI-TOF-MS and MALDI-FTICR-MS measurements respectively. This shows that the variation in the MALDI-FTICR-MS measurements is slightly lower than in the MALDI-TOF-MS measurements.

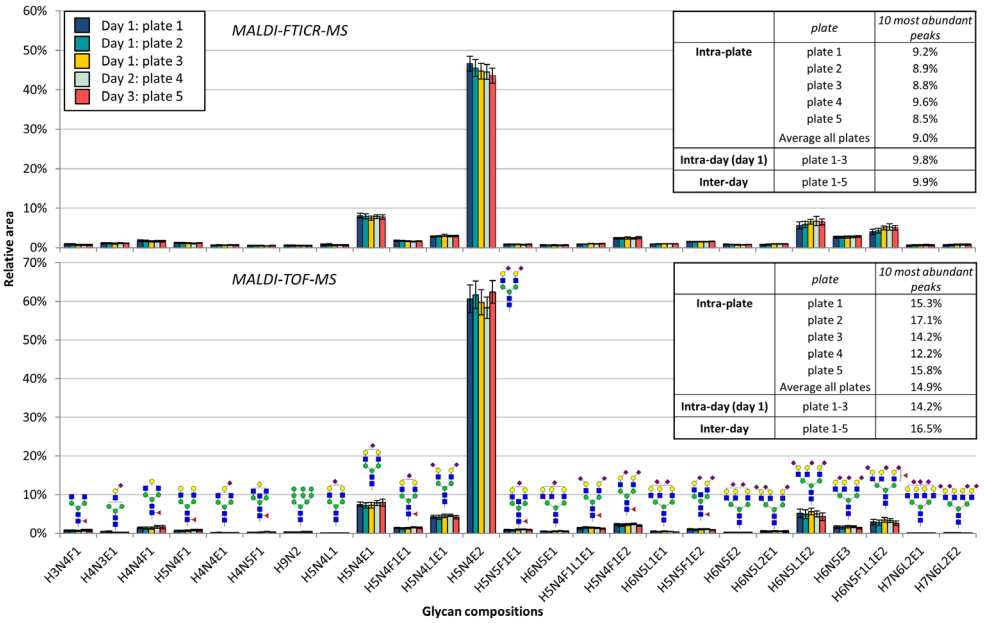


	plate	10 most abundant peaks
Intra-plate	plate 1	9.2%
	plate 2	8.9%
	plate 3	8.8%
	plate 4	9.6%
	plate 5	8.5%
	Average all plates	9.0%
Intra-day (day 1)	plate 1-3	9.8%
Inter-day	plate 1-5	9.9%

	plate	10 most abundant peaks
Intra-plate	plate 1	15.3%
	plate 2	17.1%
	plate 3	14.2%
	plate 4	12.2%
	plate 5	15.8%
	Average all plates	14.9%
Intra-day (day 1)	plate 1-3	14.2%
Inter-day	plate 1-5	16.5%

FTICR-MS measurements for the 25 most abundant glycan signals in MALDI-FTICR-MS. On day 1 the released glycan sample plate was processed three times (resulting in plate 1-3) and on day 2 and day 3 the same sample plate was processed once (respectively resulting in plate 4 and plate 5). Error bars show standard deviation. H = hexose, N = *N*-acetylhexosamine, F = deoxyhexose (fucose), L = lactonized *N*-acetylneuraminic acid (α2,3-linked), E = ethyl esterified *N*-acetylneuraminic acid (α2,6-linked).

Conclusion

The aim of the present research was to further develop and improve our previous plasma glycomics protocol, which includes linkage specific sialic acid derivatization, HILIC purification and MALDI measurement. In this study the occurrence of side-reactions during derivatization was successfully prevented and a newly automated HILIC purification method was included. Moreover, the measurements were performed on a MALDI-FTICR-MS instrument, providing ultrahigh resolution spectra with resulting higher dynamic range and precision, which were proven to be beneficial for the correct annotation and relative quantification of the observed signals. These changes were made without affecting the repeatability of the protocol. Taken together, the improvements in the methods provide a time-efficient and repeatable protocol with higher confidence identifications for glycosylation analysis in clinical applications.

Acknowledgements

The society "Genootschap Keukenhof" (Lisse, The Netherlands) financially supported parts of this research performed at the LUMC to further endorse the development of a blood-based test for early detection of cancer. No grant number applicable. The authors thank Jan Nouta and Jennifer Heikamp for their assistance with the automated pipetting platform and Guinevere S. M. Kammeijer for her valuable contribution to the glycan annotation.

References

(1) Aebersold, R.; Mann, M. Mass-Spectrometric Exploration of Proteome Structure and Function. Nature 2016, 537 (7620), 347–355. https://doi.org/10.1038/nature19949.

(2) Dell, A.; Morris, H. R. Glycoprotein Structure Determination by Mass Spectrometry. Science. March 23, 2001, pp 2351–2356. https://doi.org/10.1126/science.1058890.

(3) Cummings, R. D. The Repertoire of Glycan Determinants in the Human Glycome. Mol. Biosyst. 2009, 5 (10), 1087. https://doi.org/10.1039/b907931a.

(4) Adamczyk, B.; Tharmalingam, T.; Rudd, P. M. Glycans as Cancer Biomarkers. Biochim. Biophys. Acta - Gen. Subj. 2012, 1820 (9), 1347–1353. https://doi.org/10.1016/j.bbagen.2011.12.001.

(5) Guillard, M.; Morava, E.; Van Delft, F. L.; Hague, R.; Körner, C.; Adamowicz, M.; Wevers, R. A.; Lefeber, D. J. Plasma N-Glycan Profiling by Mass Spectrometry for Congenital Disorders of Glycosylation Type II. Clin. Chem. 2011, 57 (4), 593–602. https://doi.org/10.1373/clinchem.2010.153635.

(6) Smith, L. M.; Kelleher, N. L. Proteoform: A Single Term Describing Protein Complexity. Nat. Methods 2013, 10 (3), 186–187. https://doi.org/10.1038/nmeth.2369.

(7) Doerr, A. Glycoproteomics. Nat. Methods 2012, 9 (1), 36–36. https://doi.org/10.1038/nmeth.1821.

(8) van der Burgt, Y. E. M.; Cobbaert, C. M. Proteoform Analysis to Fulfill Unmet Clinical Needs and Reach Global Standardization of Protein Measurands in Clinical Chemistry Proteomics. Clin. Lab. Med. 2018, 38 (3), 487–497. https://doi.org/10.1016/j.cll.2018.05.001.

(9) Shubhakar, A.; Reiding, K. R.; Gardner, R. a.; Spencer, D. I. R.; Fernandes, D. L.; Wuhrer, M. High-Throughput Analysis and Automation for Glycomics Studies. Chromatographia 2015, 78 (5–6), 321–333. https://doi.org/10.1007/s10337-014-2803-9.

(10) van den Broek, I.; Nouta, J.; Razavi, M.; Yip, R.; Bladergroen, M. R.; Romijn, F. P. H. T. M.; Smit, N. P. M.; Drews, O.; Paape, R.; Suckau, D.; et al. Quantification of Serum Apolipoproteins A-I and B-100 in Clinical Samples Using an Automated SISCAPA-MALDI-TOF-MS Workflow. Methods 2015, 81, 74–85. https://doi.org/10.1016/j.ymeth.2015.03.001.

(11) Fu, Q.; Kowalski, M. P.; Mastali, M.; Parker, S. J.; Sobhani, K.; Van Den Broek, I.; Hunter, C. L.; Van Eyk, J. E. Highly Reproducible Automated Proteomics Sample Preparation Workflow for Quantitative Mass Spectrometry. J. Proteome Res. 2018, 17 (1), 420–428. https://doi.org/10.1021/acs.jproteome.7b00623.

(12) Bladergroen, M. R.; Reiding, K. R.; Hipgrave Ederveen, A. L.; Vreeker, G. C. M.; Clerc, F.; Holst, S.; Bondt, A.; Wuhrer, M.; van der Burgt, Y. E. M. Automation of High-Throughput Mass Spectrometry-Based Plasma N -Glycome Analysis with Linkage-Specific Sialic Acid Esterification. J. Proteome Res. 2015, 14 (9), 4080–4086. https://doi.org/10.1021/acs.jproteome.5b00538.

(13) Huffman, J. E.; Pučić-Baković, M.; Klarić, L.; Hennig, R.; Selman, M. H. J.; Vučković, F.; Novokmet, M.; Krištić, J.; Borowiak, M.; Muth, T.; et al. Comparative Performance of Four Methods for High-Throughput Glycosylation Analysis of Immunoglobulin G in Genetic and Epidemiological Research. Mol. Cell. Proteomics 2014, 13 (6), 1598–1610. https://doi.org/10.1074/mcp.M113.037465.

(14) Váradi, C.; Lew, C.; Guttman, A. Rapid Magnetic Bead Based Sample Preparation for Automated and High Throughput N-Glycan Analysis of Therapeutic Antibodies. Anal. Chem. 2014, 86 (12), 5682–5687. https://doi.org/10.1021/ac501573g.

(15) Szigeti, M.; Lew, C.; Roby, K.; Guttman, A. Fully Automated Sample Preparation for Ultrafast N-Glycosylation Analysis of Antibody Therapeutics. J. Lab. Autom. 2016, 21 (2), 281–286. https://doi.org/10.1177/2211068215608767.

(16) Stöckmann, H.; Duke, R. M.; Millán Martín, S.; Rudd, P. M. Ultrahigh Throughput, Ultrafiltration-Based N-Glycomics Platform for Ultraperformance Liquid Chromatography (ULTRA3). Anal. Chem. 2015, 87 (16), 8316–8322. https://doi.org/10.1021/acs.analchem.5b01463.

(17) Yang, S.; Clark, D.; Liu, Y.; Li, S.; Zhang, H. High-Throughput Analysis of N-Glycans Using AutoTip via Glycoprotein Immobilization. Sci. Rep. 2017, 7 (1), 1–11. https://doi.org/10.1038/s41598-017-10487-8.

(18) Adamczyk, B.; Stöckmann, H.; O'Flaherty, R.; Karlsson, N. G.; Rudd, P. M. High-Throughput Analysis of the Plasma N-Glycome by UHPLC. In High-Throughput Glycomics and Glycoproteomics; Lauc, G., Wuhrer, M., Eds.; Humana Press, New York, NY: New York, 2017; pp 97–108. https://doi.org/10.1007/978-1-4939-6493-2_8.

(19) Doherty, M.; Bones, J.; McLoughlin, N.; Telford, J. E.; Harmon, B.; DeFelippis, M. R.; Rudd, P. M. An Automated Robotic Platform for Rapid Profiling Oligosaccharide Analysis of Monoclonal Antibodies Directly from Cell Culture. Anal. Biochem. 2013, 442 (1), 10–18. https://doi.org/10.1016/j.ab.2013.07.005.

(20) Stöckmann, H.; Adamczyk, B.; Hayes, J.; Rudd, P. M. Automated, High-Throughput IgG-Antibody Glycoprofiling

Platform. Anal. Chem. 2013, 85 (18), 8841–8849. https://doi.org/10.1021/ac402068r.

(21) Baković, M. P.; Selman, M. H. J.; Hoffmann, M.; Rudan, I.; Campbell, H.; Deelder, A. M.; Lauc, G.; Wuhrer, M. High-Throughput IgG Fc N-Glycosylation Profiling by Mass Spectrometry of Glycopeptides. J. Proteome Res. 2013, 12 (2), 821–831. https://doi.org/10.1021/pr300887z.

(22) Nishimura, S. I.; Niikura, K.; Kurogochi, M.; Matsushita, T.; Fumoto, M.; Hinou, H.; Kamitani, R.; Nakagawa, H.; Deguchi, K.; Miura, N.; et al. High-Throughput Protein Glycomics: Combined Use of Chemoselective Glycoblotting and MALDI-TOF/TOF Mass Spectrometry. Angew. Chemie - Int. Ed. 2004, 44 (1), 91–96. https://doi.org/10.1002/anie.200461685.

(23) Ruhaak, L. R.; Huhn, C.; Waterreus, W. J.; De Boer, A. R.; Neusüss, C.; Hokke, C. H.; Deelder, A. M.; Wuhrer, M. Hydrophilic Interaction Chromatography-Based High-Throughput Sample Preparation Method for N-Glycan Analysis from Total Human Plasma Glycoproteins. Anal. Chem. 2008, 80 (15), 6119–6126. https://doi.org/10.1021/ac800630x.

(24) Reiding, K. R.; Blank, D.; Kuijper, D. M.; Deelder, A. M.; Wuhrer, M. High-Throughput Profiling of Protein N-Glycosylation by MALDI-TOF-MS Employing Linkage-Specific Sialic Acid Esterification. Anal. Chem. 2014, 86 (12), 5784–5793. https://doi.org/10.1021/ac500335t.

(25) Reiding, K. R.; Vreeker, G. C. M.; Bondt, A.; Bladergroen, M. R.; Hazes, J. M. W.; van der Burgt, Y. E. M.; Wuhrer, M.; Dolhain, R. J. E. M. Serum Protein N-Glycosylation Changes with Rheumatoid Arthritis Disease Activity during and after Pregnancy. Front. Med. 2018, 4 (241), 1–11. https://doi.org/10.3389/fmed.2017.00241.

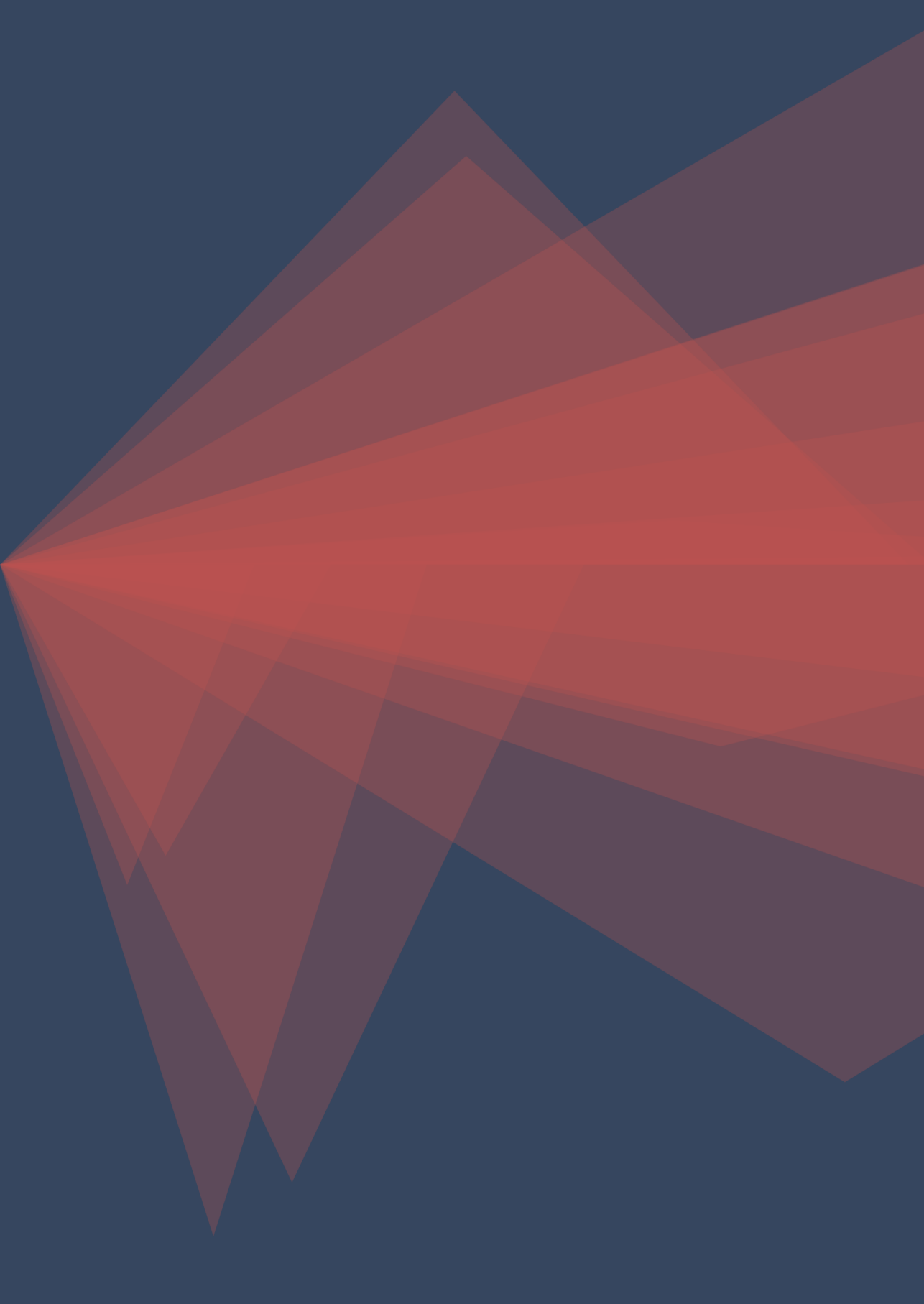
(26) Selman, M. H. J.; Hemayatkar, M.; Deelder, A. M. A. M.; Wuhrer, M. Cotton HILIC SPE Microtips for Microscale Purification and Enrichment of Glycans and Glycopeptides. Anal. Chem. 2011, 83 (7), 2492–2499. https://doi.org/10.1021/ac1027116.

(27) Carr, S. A.; Abbatiello, S. E.; Ackermann, B. L.; Borchers, C.; Domon, B.; Deutsch, E. W.; Grant, R. P.; Hoofnagle, A. N.; Hüttenhain, R.; Koomen, J. M.; et al. Targeted Peptide Measurements in Biology and Medicine: Best Practices for Mass Spectrometry-Based Assay Development Using a Fit-for-Purpose Approach. Mol. Cell. Proteomics 2014, 13 (3), 907–917. https://doi.org/10.1074/mcp.M113.036095.

(28) Lim, M. D.; Dickherber, A.; Compton, C. C. Before You Analyze a Human Specimen, Think Quality, Variability, and Bias. Anal. Chem. 2011, 83 (1), 8–13. https://doi.org/10.1021/ac1018974.

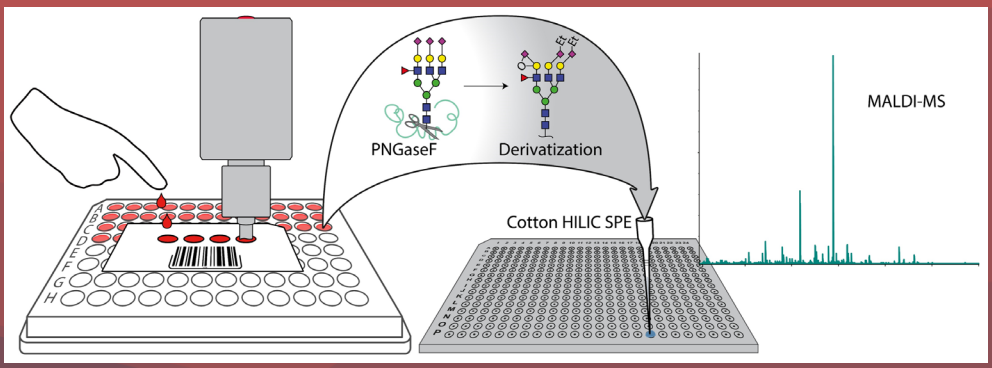
(29) Reiding, K. R.; Lonardi, E.; Hipgrave Ederveen, A. L.; Wuhrer, M. Ethyl Esterification for MALDI-MS Analysis of Protein Glycosylation. Methods Mol. Biol. 2016, 1394, 151–162. https://doi.org/10.1007/978-1-4939-3341-9_11.

(30) Nicolardi, S.; Bogdanov, B.; Deelder, A. M.; Palmblad, M.; Van Der Burgt, Y. E. M. Developments in FTICR-MS and Its Potential for Body Fluid Signatures. Int. J. Mol. Sci. 2015, 16 (11), 27133–27144. https://doi.org/10.3390/ijms161126012.



CHAPTER 3

Dried blood spot *N*-glycome analysis by MALDI mass spectrometry



Based on:

Vreeker, G.C.M.[1,2]*., Bladergroen, M.R.*[1]*, Nicolardi, S.*[1]*, Mesker, W.E.*[1]*, Tollenaar, R.A.E.M.*[1]*, van der Burgt, Y.E.M.*[1,3]*, Wuhrer, M.*[1]*: Dried blood spot N-glycome analysis by MALDI mass spectrometry. Talanta. 205, 120104 (2019). doi:10.1016/j.talanta.2019.06.104. Supplemental info available.*

1 Center for Proteomics and Metabolomics, Leiden University Medical Center, The Netherlands

2 Department of Surgery, Leiden University Medical Center, The Netherlands

3 Department of Clinical Chemistry and Laboratory Medicine, Leiden University Medical Center, The Netherlands

Abstract

Body fluid *N*-glycome analysis as well as glyco-proteoform profiling of existing protein biomarkers potentially provides a stratification layer additional to quantitative, diagnostic protein levels. For clinical omics applications, the collection of a dried blood spot (DBS) is increasingly pursued as an alternative to sampling milliliters of peripheral blood. Here we evaluate DBS cards as a blood collection strategy for protein *N*-glycosylation analysis aiming for high-throughput clinical applications. A protocol for facile *N*-glycosylation profiling from DBS is developed that includes sialic acid linkage differentiation. This protocol is based on a previously established total plasma *N*-glycome mass spectrometry (MS) method, with adjustments for the analysis of DBS specimens. After DBS-punching and protein solubilization *N*-glycans are released, followed by chemical derivatization of sialic acids and MS-measurement of *N*-glycan profiles. With this method, more than 80 different glycan structures are identified from a DBS, with RSDs below 10% for the ten most abundant glycans. *N*-glycan profiles of finger-tip blood and venous blood are compared and short-term stability of DBS is demonstrated. This method for fast *N*-glycosylation profiling of DBS provides a minimally invasive alternative to conventional serum and plasma protein *N*-glycosylation workflows. With simplified blood sampling this DBS approach has vast potential for clinical glycomics applications.

Introduction

Total plasma or serum *N*-glycosylation analysis has been widely performed to study human protein glycosylation in general, and aberrant disease-related glycosylation profiles or alterations specifically.[1–9] Various strategies can be pursued to obtain *N*-glycome readouts from (glyco)proteins present in the circulation[10], however for large-scale evaluations of clinical cohorts throughput numbers become essential. Commonly blood samples (for sequential plasma- or serum-based glycome analysis) are collected by antecubital venipuncture, which needs to be performed by a skilled person. An alternative and less invasive micro blood-sampling procedure to obtain specimens is through collection on a filter paper with a simple finger prick. Additional advantages of such a dried blood spot (DBS) are the ease of transportation and storage of the samples.[11,12] Consequently, DBS collection holds great potential for population-wide biomarker screening programs.[13] Hitherto DBS collections are known from newborn screening programs for metabolic diseases, but dried blood spots are also convenient for pharmacokinetic studies, diagnosis of infectious diseases and monitoring drug therapies.[14–17] Routine measurement of nucleic acids, small molecules and lipids from DBS specimens is pursued in clinical chemistry laboratories and for such determinations mass spectrometry (MS) is the method of choice.[18] Most commonly, MS-analysis is performed after punching a part of the DBS card, however also paper spray- and liquid extraction strategies have been reported that leave the card intact.[19,20]

DBS protein analysis has become feasible as a result of the developments in MS-based proteomics technologies.[21,22] Examples of MS-based protein analysis obtained from DBS are the measurement of glycated and acetylated hemoglobin[23–25] and quantification of ceruloplasmin for newborn screening of Wilson disease.[26]. Current bottom-up proteomics strategies allow for the quantification of 97 proteins.[27] With regard to protein glycosylation analysis, up to now only a couple of examples have demonstrated the analysis of protein glycosylation from DBS, despite its vast biomarker potential.[28] The first one applied enzymatic release of *N*-glycans with subsequent analysis using liquid chromatography(LC)-porous graphitized carbon (PGC)- time-of-flight(TOF)-MS, resulting in the identification and relative quantification of 44 glycan compositions from DBS samples.[29] The second example of DBS protein glycosylation analysis followed an *N*-glycopeptide approach, resulting in the identification of 41 *N*-glycopeptides from 16 major *N*-glycoproteins.[11]

In the current study, we apply a novel high-throughput glycomics method that allows profiling of more than 80 different *N*-glycans from DBS samples with linkage-specific sialic acid analysis using matrix-assisted laser desorption/ionization (MALDI) MS. Furthermore, aiming for a simplified blood collection strategy,

potential stress during sample transportation and storage needs to be considered.[30] Therefore, short-term stability of the *N*-glycome from a DBS will be evaluated at different storage conditions.

Experimental section

Chemicals and enzymes

Milli-Q water was produced with a Q-Gard 2 system at ≥18 MΩ (Millipore). Recombinant peptide-*N*-glycosidase F (PNGase F) was purchased from Roche Diagnostics and potassium dihydrogen phosphate (KH_2PO_4), 50% sodium hydroxide (NaOH), 85% phosphoric acid (H_3PO_4), disodium hydrogen phosphate dihydrate ($Na_2HPO_4 \times 2\ H_2O$), sodium chloride (NaCl), super-DHB (9:1 mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid, sDHB), nonidet P-40 substitute (NP-40) and 1-hydroxybenzotriazole 97% (HOBt) were obtained from Sigma-Aldrich. Potassium hydroxide (KOH), trifluoroacetic acid (TFA), analytical grade ethanol and sodium dodecyl sulfate (SDS), and were obtained from Merck. High performance LC-grade acetonitrile (ACN) was obtained from Biosolve and 1-ethyl-3-(3-(dimethylamino)propyl)carbodiimide (EDC) hydrochloride was purchased from Fluorochem.

Samples

The blood- and plasma samples were collected on Whatman FTA DMPK-C DBS cards from healthy volunteers and fully anonymized. For the stability study at -80 °C, from each volunteer A and B fourteen DBS samples were collected. For the stability study at room temperature sixteen DBS samples were collected from volunteer C. The DBS samples from volunteers A, B and C were, due to the number of blood drops needed for method validation, spotted with a glass pipette from collected EDTA-blood tubes (directly after collection) and will therefore be referred to as DBS (venous). For the stability study at 37 °C eight DBS samples from volunteer D were collected via a finger prick and will therefore be referred to as DBS (finger). The spots from volunteer D were made by washing the hands of the volunteer with water and pressing a lancet (Accu-Chek Safe T-pro plus, 2.3 mm deep) on the fingertip of the volunteer. The finger was gently massaged and the drops of blood were collected. All blood spots were left to dry for 24 hours at room temperature, before they were stored or before further processing. Per DBS two disks of approximately 0.25 cm^2 were punched with an inhouse developed electric puncher.

Enzymatic *N*-glycan release

A punched-out filter paper disk with the DBS was added to 40 µL 2% SDS in an 2 mL Eppendorf tube or microtitration plate (MTP, Nunc 96 wells plate, Thermo Fisher) and incubated for 10 minutes at 60 °C. Subsequently 41 µL of releasing mixture (20 µL acidic 5x PBS (pH 5.6), 20 µL 4% NP-40 and 1 µL PNGase F) was added. The slightly acidic condition in the release reaction was used to rapidly hydrolyze the

glycosylamines which are formed after PNGase F release, to prevent subsequent side reactions of these glycosylamines.[31] The sample was incubated overnight at 37 °C. The samples were directly used for ethyl esterification or stored at -20 °C for later use.

Ethyl esterification

Linkage-specific derivatization of sialic acids through ethylation of α2,6-linked sialic acid residues and via lactone formation of α2,3-linked sialic acids was performed as described before.[32] In short, 2 μL of released glycan sample was added to 20 μL ethyl esterification reagent (0.25 M EDC and 0.25 M HOBt in ethanol). The samples were incubated for one hour at 37 °C. Afterwards 22 μL of ACN was added.

Cotton HILIC purification

Purification of the glycans was performed with cotton hydrophilic interaction liquid chromatography HILIC tips, as described previously.[32,33] In short, 3 mm cotton thread (approximately 180 μg, Pipoos, Utrecht, The Netherlands) was positioned at the bottom of 20 μL tips. The cotton was prewetted with three times 20 μL MQ and subsequently three times conditioned with 85% ACN. The sample was loaded on the cotton by pipetting the sample twenty times up and down. The cotton was washed with first three times 20 μL 85% ACN with 1% TFA and second three times 20 μL 85% ACN. The glycans were eluted from the cotton by pipetting five times up and down in 10 μL MQ. For the analysis 1 μL of sDHB matrix (5 mg mL-1 in 99% ACN with 1 mM NaOH) was spotted with 2 μL of purified glycan sample.

MALDI-TOF and MALDI-FTICR mass measurements

Initially, our standard protocol included MALDI-TOF-MS analysis and such measurements were applied. In the progress of this project, MALDI-Fourier-transform ion cyclotron resonance (FTICR) MS analysis became available and was used for parts of the study.[31] MALDI-TOF-MS and MALDI-FTICR-MS spectra were recorded as described before.[31] In short, for the MALDI-TOF-MS measurements (*m*/*z*-range 1000-5000) an UltrafleXtreme mass spectrometer (Bruker Daltonics) was used in reflectron positive mode with a Smartbeam-II laser. The system was controlled by flexControl software (version 3.4 Build 135). At a frequency of 1000 Hz, 10000 laser shots were shot in a random walking pattern, with 200 shots per raster spot. The MALDI-FTICR-MS spectra were recorded from *m*/*z* 1011 to 5000 on a 15T SolariX XR FTICR mass spectrometer (Bruker Daltonics). The system was operated by ftmsControl (version 2.1.0) and each spectrum consisted of ten scans with 200 laser shots at each raster.

Comparison DBS (venous and from finger) with dried plasma spots (DPS)
From four anonymous volunteers a tube (EDTA) of blood and four DBS (finger) were collected. A drop of blood from the tube was spotted onto each of the marked areas of an additional DBS card. The tubes were subsequently centrifuged at 1000 g for 10 minutes and additionally four plasma spots per volunteer were created. All spots were left to dry for 24 hours, where after the derivatization, purification and measurements were performed.

Short-term stability study
The short-term stability of the glycans in DBS was evaluated for different storage temperatures. Twelve out of fourteen spots of volunteer A and B were stored at -80 °C for two, four and six weeks (per four spots). From the other two spots per volunteer the glycans were directly released. Also, for volunteer C twelve spots were stored at room temperature for two, four and six weeks (per four spots). From the other four spots the glycans were also directly released. Additionally, four spots of volunteer D were stored at 37 °C for two weeks and from four spots of the same volunteer the glycans were again directly released. At the given time points, the glycans were released from the cards that were stored in the freezer. After the release of the glycans from the cards at the last timepoint in each series the derivatization, purification and measurements were performed for all releases from all timepoints.

MS-data preprocessing
The obtained spectra were transformed in text format (x,y) using flexAnalysis version 3.4 and DataAnalysis version 3.1 (both Bruker Daltonics) for respectively the MALDI-TOF-MS and MALDI-FTICR-MS spectra. The resulting text files were calibrated using MassyTools (version 0.1.8.1). In Supporting Information Table S-1 the list of calibration masses is shown. The spectrum curation and analyte curation was performed according to the criteria described previously.[31] For the curation of analytes a list of 134 glycan compositions including 19 dummy signals (Supporting Information Table S-2) was extracted from each spectrum. Analytes were only included in a final extraction of the data if the criteria on S/N ratio, isotopic pattern quality and ppm-error were met. The final analyte lists for MALDI-TOF-MS and MALDI-FTICR-MS spectra can be found in Supporting Information Table S-3. After the final extraction of the signals, the extracted peak-areas are normalized for each individual spectrum to the sum of all signal areas.

Results

DBS sample preparation for *N*-glycome profiles

A protocol was developed that included punching of a disk, solubilization of dried blood with subsequent *N*-glycan release from blood glycoproteins, linkage-specific derivatization of the sialic acids to differentiate between α2,3- and α2,6-linkages, HILIC purification of the released *N*-glycans and MALDI-MS measurement.

The DBS samples were punched out and (glyco)proteins were solubilized in a buffer containing PNGase F to release the *N*-glycans. Thus obtained glycans were subsequently derivatized, purified with cotton HILIC solid phase extraction(SPE) tips, spotted onto a MALDI-target plate and measured with MALDI-MS. All glycans were observed as sodiated species $[M+Na]^+$. MALDI-TOF spectra of *N*-glycans released from DBS were similar with previously reported MALDI-TOF *N*-glycome spectra of serum and plasma (spectra not shown).[4,32,34] In Fig. 1, a typical MALDI-FTICR-MS spectrum of the DBS *N*-glycome is shown, which resembled the earlier reported MALDI-FTICR spectra of plasma protein *N*-glycans.[31] Note that all glycan abundances are reported as relative numbers that are derived after normalization based on the summed glycan intensities. This means that if some signals would increase, other signals will decrease.

The positional and linkage assignments of the glycans were based on previously reported identifications[31,32], except for the linkage of the *N*-acetylneuraminic acid moieties that is inherently determined by the mass shift after derivatization. The reaction scheme of this derivatization is shown in Supporting Information Fig. S-1. From the MALDI-TOF-MS and MALDI-FTICR-MS spectra obtained from DBS respectively 82 and 84 glycan signals were extracted for relative quantification (Supporting Information Table S-3). For both MALDI-TOF-MS and MALDI-FTICR-MS measurements the relative standard deviation (RSD) was below 10% for the ten most abundant glycan signals (accounting for >75% of total analyte area).

Comparison between DBS (venous/from finger) with dried plasma spots (DPS)

The first step in our study concerned the evaluation of overlap as well as potential differences between a dried blood spot and a dried plasma spot. The results are summarized in Fig. 2, and it is demonstrated that glycan profiles obtained from the different samples are highly similar. No differences are observed between the glycan profiles from DBS obtained from a finger prick and from venipuncture, or at least differences are not larger than the technical variation of 10%. In the case of DPS, the glycan profiles slightly differed for the signals annotated as H3N4F1, H4N4F1 and H5N4F1, with a higher signal in the DPS samples and lower corresponding major biantennary glycan H5N4E2

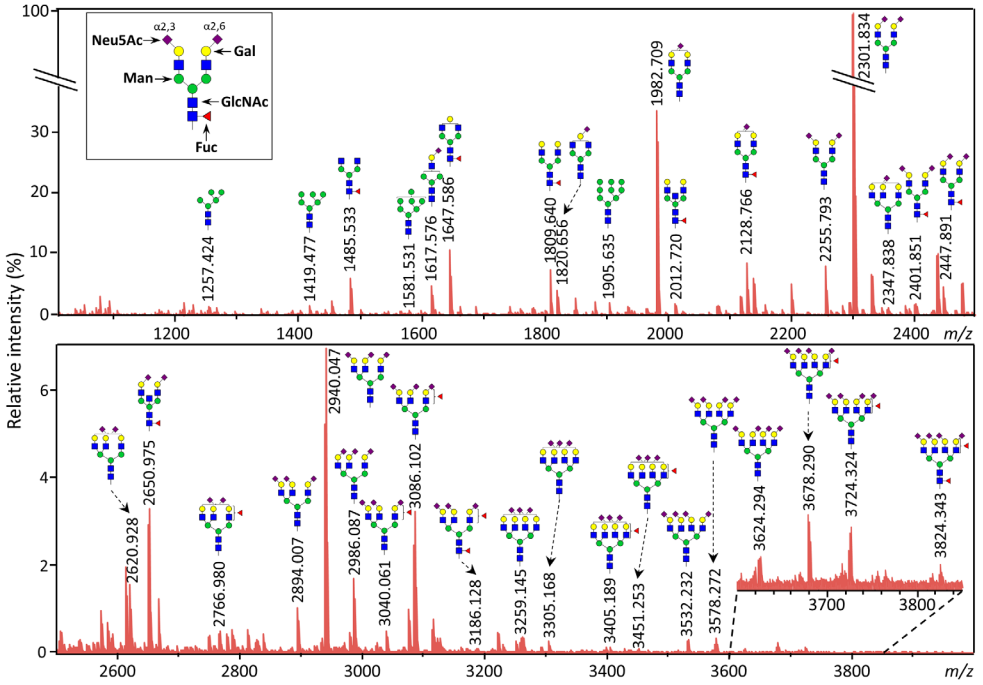


Figure 1. MALDI-FTICR mass spectrum of released *N*-glycome from dried blood spots. Sialic acids are derivatized to allow differentiation between α2,3- and α2,6-linked residues and all glycans are assigned as [M+Na]+. The most intense signals were annotated in the figure, the total list of 84 *N*-glycan compositions which were identified can be found in Supporting Information Table S-3.

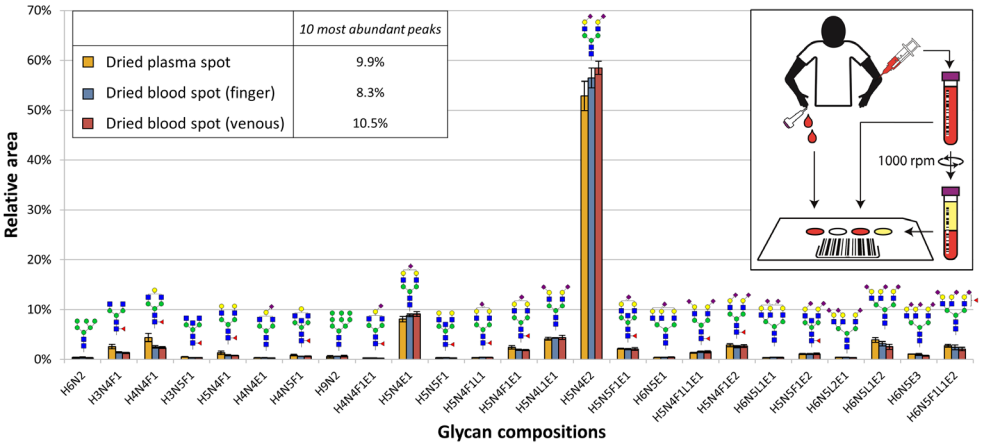


Figure 2. Comparison of MALDI-TOF-MS glycan profiles obtained from plasma, venous blood, and finger-tip blood of a single person. The 25 most abundant glycan compositions are shown. Error bars show standard deviation. H = hexose, N = *N*-acetylhexosamine, F = deoxyhexose (fucose), L = lactonized *N*-acetylneuraminic acid (α2,3-linked), E = ethyl esterified *N*-acetylneuraminic acid (α2,6-linked).

DBS storage conditions

Multiple storage temperatures and time periods were evaluated for the storage of the DBS on the filter paper cards. For the storage at -80 °C and room temperature four time points were measured with a maximum of six weeks storage. The obtained glycan profiles from the spots are constant over time at both storage temperatures, this is shown in Supporting Information Fig. S-2 and S-3. Interestingly, the DBS samples that were stored at 37 °C show similar profiles to the profiles from samples that were directly processed (Fig. 3).

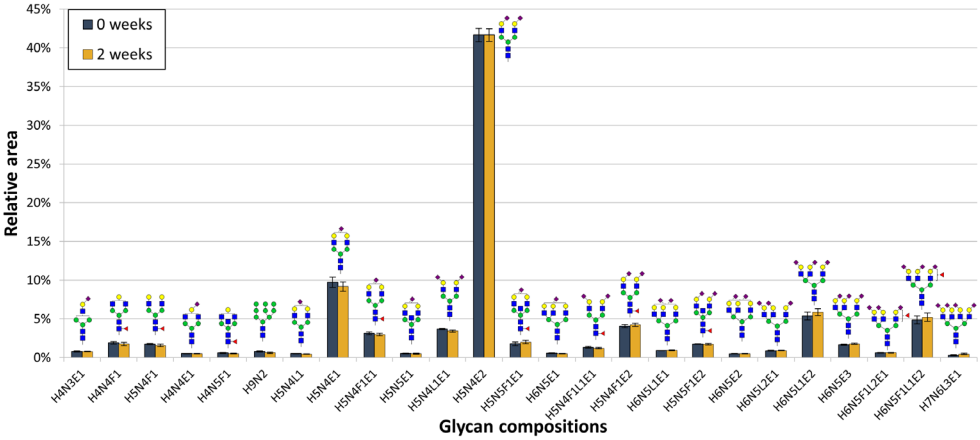


Figure 3. Comparison of MALDI-FTICR-MS glycan profiles from DBS which were stored at 37 °C for two weeks and DBS which, after drying, were directly processed. Error bars show standard deviation. H = hexose, N = *N*-acetylhexosamine, F = deoxyhexose (fucose), L = lactonized *N*-acetylneuraminic acid (α2,3-linked), E = ethyl esterified *N*-acetylneuraminic acid (α2,6-linked).

Discussion

In this study, we present a workflow for *N*-glycan profiling from a DBS with mass measurement precision suitable for clinical cohort analysis. As a starting point our in-house developed plasma glycomics analysis protocol was used.[31] Several changes compared to regular plasma samples were required in order to obtain glycan profiles from a DBS, with the largest adjustment punching out a disk from the DBS. Before punching, the device reads the 2D-barcode of the DBS card ensuring correct track-and-trace of the sample. Such barcode reading prior to punching in combination with fully automated liquid handling of all pipetting step allows screening of large numbers of DBS cards in a standardized way, with throughput numbers up to six MTPs per day (i.e. 576 samples). After the DBS disk is punched and captured in a round-bottom MTP, proteins are solubilized (extracted) from the filter paper using a 2% SDS solution that is identical to the one in the initial plasma protocol. A smaller adjustment involved the total volume of the glycan release reaction that was increased to ensure the disks of the DBS were completely covered with liquid.
MALDI-MS analysis was performed on an FTICR-system for confident glycan identification. The short-term stability experiments were carried out on TOF-MS, which demonstrates that mass measurements are not necessarily limited to an ultrahigh resolution instrument, however it was noted that this slightly decreased the number of identifications. The repertoire of *N*-glycans found in the spectra is highly similar to that obtained from plasma samples using the recently established workflow and exhibits similar repeatability.[31] Furthermore, since quantification of glycans occurs in a relative manner, the analyses are suited for the inherently varying volumes of blood drops on the filter card.

The differences between DPS and DBS glycan profiles were found to be minor (Figure 2). These minor differences comprise three specific immunoglobulin glycans, namely H3N4F1, H4N4F1 and H5N4F1, which is in excellent agreement with a previous report where IgG-specific glycopeptides were lower in abundance in a DBS compared to plasma.[11] Overall it can be concluded that most of the released glycans originate from plasma proteins and that the structural assignment for plasma glycans can be used for the identification of signals in the DBS spectra. Also, the observed signals in spectra from DBS met the criteria for analyte curation with the same list of compositions as the plasma samples.

Glycans are known for their long-term chemical stability. An extreme example hereof is the detection of specific parasite epitopes in ancient mummies.[35] With regard to DBS storage, the stability of glycopeptides was previously evaluated over a period of 180 days and it was found that the summed intensity of 32 *N*-glycopeptides from 14 abundant *N*-glycoproteins did not change.[11] In the current

study, the stability of released *N*-glycans was monitored over a 6-week time-period. Glycosylation profiles did not change after storage at different temperatures (-80 °C, room temperature and 37 °C) over various periods. The confirmation that the storage temperature does not affect the glycan profiles makes it easier to set up a protocol for self-sampling of DBS from patients who have difficulty traveling to a doctor's office or who need regular testing over a longer period.[36,37] Thus obtained DBS sample collections can be shipped to a central facility for *N*-glycome analysis.

In conclusion, this study reports the development and performance of a method for the analysis of the total *N*-glycome from DBS. We have followed up on previous efforts to profile *N*-glycans from DBS and identified and relatively quantified a larger number of glycans. The short-term stability of glycan profiles over time was evaluated for different storage conditions and these results indicated a great potential for facile shipment and storage of DBS cards. The availability of DBS will simplify access to diagnostic tests on the basis of *N*-glycomic analyses since samples can be collected by an individual wherever they are, including at home. We foresee that DBS *N*-glycomic approaches similar to the one outlined here have great potential for clinical biomarker development and implementation.

Acknowledgements

This work was supported by the society "Genootschap ter ondersteuning van de vroege opsporing van kanker" (Lisse, The Netherlands) to further endorse the development of a blood-based test for early detection of cancer (no grant number applicable). The authors thank the anonymous volunteers for their donations and MDs Kiki Vangangelt and Randa Sawires for their assistance in the sample collection.

References

(1) Meany, D. L.; Zhang, Z.; Sokoll, L. J.; Zhang, H.; Chan, D. W. Glycoproteomics for Prostate Cancer Detection: Changes in Serum PSA Glycosylation Patterns. J. Proteome Res. 2009, 8 (2), 613–619. https://doi.org/10.1021/pr8007539.

(2) de Vroome, S. W.; Holst, S.; Girondo, M. R.; van der Burgt, Y. E. M.; Mesker, W. E.; Tollenaar, R. A. E. M.; Wuhrer, M. Serum N-Glycome Alterations in Colorectal Cancer Associate with Survival. Oncotarget 2018, 9 (55), 30610–30623. https://doi.org/10.18632/oncotarget.25753.

(3) Juszczak, A.; Pavić, T.; Vučković, F.; Bennett, A. J.; Shah, N.; Medvidović, E. P.; Groves, C. J.; Šekerija, M.; Chandler, K.; Burrows, C.; et al. Plasma Fucosylated Glycans and C-Reactive Protein As Biomarkers of HNF1A-MODY in Young Adult-Onset Nonautoimmune Diabetes. Diabetes Care 2018, No. 1. https://doi.org/10.2337/dc18-0422.

(4) Reiding, K. R.; Vreeker, G. C. M.; Bondt, A.; Bladergroen, M. R.; Hazes, J. M. W.; van der Burgt, Y. E. M.; Wuhrer, M.; Dolhain, R. J. E. M. Serum Protein N-Glycosylation Changes with Rheumatoid Arthritis Disease Activity during and after Pregnancy. Front. Med. 2018, 4 (241), 1–11. https://doi.org/10.3389/fmed.2017.00241.

(5) Arnold, J. N.; Saldova, R.; Galligan, M. C.; Murphy, T. B.; Mimura-Kimura, Y.; Telford, J. E.; Godwin, A. K.; Rudd, P. M. Novel Glycan Biomarkers for the Detection of Lung Cancer. J. Proteome Res. 2011, 10 (4), 1755–1764. https://doi.org/10.1021/pr101034t.

(6) Kyselova, Z.; Mechref, Y.; Al Bataineh, M. M.; Dobrolecki, L. E.; Hickey, R. J.; Vinson, J.; Sweeney, C. J.; Novotny, M. V. Alterations in the Serum Glycome Due to Metastatic Prostate Cancer. J. Proteome Res. 2007, 6 (5), 1822–1832. https://doi.org/10.1021/pr060664t.

(7) Kirwan, A.; Utratna, M.; O'Dwyer, M. E.; Joshi, L.; Kilcoyne, M. Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. Biomed Res. Int. 2015, 2015, 490531. https://doi.org/10.1155/2015/490531.

(8) Drake, R. R.; Jones, E. E.; Powers, T. W.; Nyalwidhe, J. O. Altered Glycosylation in Prostate Cancer. In Advances in Cancer Research; Elsevier Inc., 2015; Vol. 126, pp 345–382. https://doi.org/10.1016/bs.acr.2014.12.001.

(9) Clerc, F.; Novokmet, M.; Dotz, V.; Reiding, K. R.; de Haan, N.; Kammeijer, G. S. M.; Dalebout, H.; Bladergroen, M. R.; Vukovic, F.; Rapp, E.; et al. Plasma N-Glycan Signatures Are Associated With Features of Inflammatory Bowel Diseases. Gastroenterology 2018, 155 (3), 829–843. https://doi.org/10.1053/j.gastro.2018.05.030.

(10) Reiding, K. R.; Bondt, A.; Hennig, R.; Gardner, R. A.; O'Flaherty, R.; Trbojević-Akmačić, I.; Shubhakar, A.; Hazes, J. M. W.; Reichl, U.; Fernandes, D. L.; et al. High-Throughput Serum N -Glycomics: Method Comparison and Application to Study Rheumatoid Arthritis and Pregnancy-Associated Changes. Mol. Cell. Proteomics 2019, 18 (1), 3–15. https://doi.org/10.1074/mcp.RA117.000454.

(11) Choi, N. Y.; Hwang, H.; Ji, E. S.; Park, G. W.; Lee, J. Y.; Lee, H. K.; Kim, J. Y.; Yoo, J. S. Direct Analysis of Site-Specific N-Glycopeptides of Serological Proteins in Dried Blood Spot Samples. Anal. Bioanal. Chem. 2017, 409 (21), 4971–4981. https://doi.org/10.1007/s00216-017-0438-z.

(12) Freeman, J. D.; Rosman, L. M.; Ratcliff, J. D.; Strickland, P. T.; Graham, D. R.; Silbergeld, E. K. State of the Science in Dried Blood Spots. Clin. Chem. 2018, 64 (4), 656–679. https://doi.org/10.1373/clinchem.2017.275966.

(13) Demirev, P. A. Dried Blood Spots: Analysis and Applications. Anal. Chem. 2013, 85 (2), 779–789. https://doi.org/10.1021/ac303205m.

(14) McDade, T. W.; Williams, S.; Snodgrass, J. J. What a Drop Can Do: Dried Blood Spots as a Minimally Invasive Method for Integrating Biomarkers into Population-Based Research. Demography 2007, 44 (4), 899–925.

(15) Spooner, N.; Lad, R.; Barfield, M. Dried Blood Spots as a Sample Collection Technique for the Determination of Pharmacokinetics in Clinical Studies: Considerations for the Validation of a Quantitative Bioanalytical Method. Anal. Chem. 2009, 81 (4), 1557–1563. https://doi.org/10.1021/ac8022839.

(16) Meesters, R. J.; Hooff, G. P. State-of-the-Art Dried Blood Spot Analysis: An Overview of Recent Advances and Future Trends. Bioanalysis 2013, 5 (17), 2187–2208. https://doi.org/10.4155/bio.13.175.

(17) Müller, M. J.; Stokes, C. S.; Volmer, D. A. Quantification of the 3α and 3β Epimers of 25-Hydroxyvitamin D 3 in Dried Blood Spots by LC-MS/MS Using Artificial Whole Blood Calibration and Chemical Derivatization. Talanta 2017, 165 (October 2016), 398–404. https://doi.org/10.1016/j.talanta.2016.12.081.

(18) Lehmann, S.; Delaby, C.; Vialaret, J.; Ducos, J.; Hirtz, C. Current and Future Use of "Dried Blood Spot" Analyses in Clinical Chemistry. Clin. Chem. Lab. Med. 2013, 51 (10), 1897–1909. https://doi.org/10.1515/cclm-2013-0228.

(19) Ferreira, C. R.; Yannell, K. E.; Jarmusch, A. K.; Pirro, V.; Ouyang, Z.; Graham Cooks, R. Ambient Ionization Mass Spectrometry for Point-of-Care Diagnostics and Other Clinical Measurements. Clin. Chem. 2016, 62 (1), 99–110. https://doi.org/10.1373/clinchem.2014.237164.

(20) Wagner, M.; Tonoli, D.; Varesio, E.; Hopfgartner, G. The Use of Mass Spectrometry to Analyze Dried Blood Spots.

Mass Spectrom. Rev. 2007, 35 (3), 361–438. https://doi.org/10.1002/mas.21441.

(21) Martin, N. J.; Cooper, H. J. Challenges and Opportunities in Mass Spectrometric Analysis of Proteins from Dried Blood Spots. Expert Rev. Proteomics 2014, 11 (6), 685–695. https://doi.org/10.1586/14789450.2014.965158.

(22) van den Broek, I.; Sobhani, K.; Van Eyk, J. E. Advances in Quantifying Apolipoproteins Using LC-MS/MS Technology: Implications for the Clinic. Expert Rev. Proteomics 2017, 14 (10), 869–880. https://doi.org/10.1080/14789450.2017.1374859.

(23) Davison, A. S.; Green, B. N.; Roberts, N. B. Fetal Hemoglobin: Assessment of Glycation and Acetylation Status by Electrospray Ionization Mass Spectrometry. Clin. Chem. Lab. Med. 2008, 46 (9), 1230–1238. https://doi.org/10.1515/CCLM.2008.257.

(24) Daniel, Y. A.; Turner, C.; Haynes, R. M.; Hunt, B. J.; Dalton, R. N. Quantification of Hemoglobin A2by Tandem Mass Spectrometry. Clin. Chem. 2007, 53 (8), 1448–1454. https://doi.org/10.1373/clinchem.2007.088682.

(25) Edwards, R. L.; Creese, A. J.; Baumert, M.; Griffiths, P.; Bunch, J.; Cooper, H. J. Hemoglobin Variant Analysis via Direct Surface Sampling of Dried Blood Spots Coupled with High-Resolution Mass Spectrometry. Anal. Chem. 2011, 83 (6), 2265–2270. https://doi.org/10.1021/ac1030804.

(26) DeWilde, A.; Sadilkova, K.; Sadilek, M.; Vasta, V.; Si, H. H. Tryptic Peptide Analysis of Ceruloplasmin in Dried Blood Spots Using Liquid Chromatography-Tandem Mass Spectrometry: Application to Newborn Screening. Clin. Chem. 2008, 54 (12), 1961–1968. https://doi.org/10.1373/clinchem.2008.111989.

(27) Chambers, A. G.; Percy, A. J.; Yang, J.; Borchers, C. H. Multiple Reaction Monitoring Enables Precise Quantification of 97 Proteins in Dried Blood Spots. Mol. Cell. Proteomics 2015, 14 (11), 3094–3104. https://doi.org/10.1074/mcp.O115.049957.

(28) Glycosciences Committee on Assessing the Importance and Impact of Technology and Glycomics. Transforming Glycoscience: A Roadmap for the Future; National Academies Press: Washington, D.C., 2012. https://doi.org/10.17226/13446.

(29) Ruhaak, L. R.; Miyamoto, S.; Kelly, K.; Lebrilla, C. B. N-Glycan Profiling of Dried Blood Spots. Anal Chem 2012, 84 (1), 396–402. https://doi.org/10.1021/ac202775t.

(30) Han, J.; Higgins, R.; Lim, M. D.; Lin, K.; Yang, J.; Borchers, C. H. Short-Term Stabilities of 21 Amino Acids in Dried Blood Spots. Clin. Chem. 2018, 64 (2), 400–402. https://doi.org/10.1373/clinchem.2017.278457.

(31) Vreeker, G. C. M.; Nicolardi, S.; Bladergroen, M. R.; van der Plas, C. J.; Mesker, W. E.; Tollenaar, R. A. E. M.; van der Burgt, Y. E. M.; Wuhrer, M. Automated Plasma Glycomics with Linkage-Specific Sialic Acid Esterification and Ultrahigh Resolution MS. Anal. Chem. 2018, 90 (20), 11955–11961. https://doi.org/10.1021/acs.analchem.8b02391.

(32) Reiding, K. R.; Blank, D.; Kuijper, D. M.; Deelder, A. M.; Wuhrer, M. High-Throughput Profiling of Protein N-Glycosylation by MALDI-TOF-MS Employing Linkage-Specific Sialic Acid Esterification. Anal. Chem. 2014, 86 (12), 5784–5793. https://doi.org/10.1021/ac500335t.

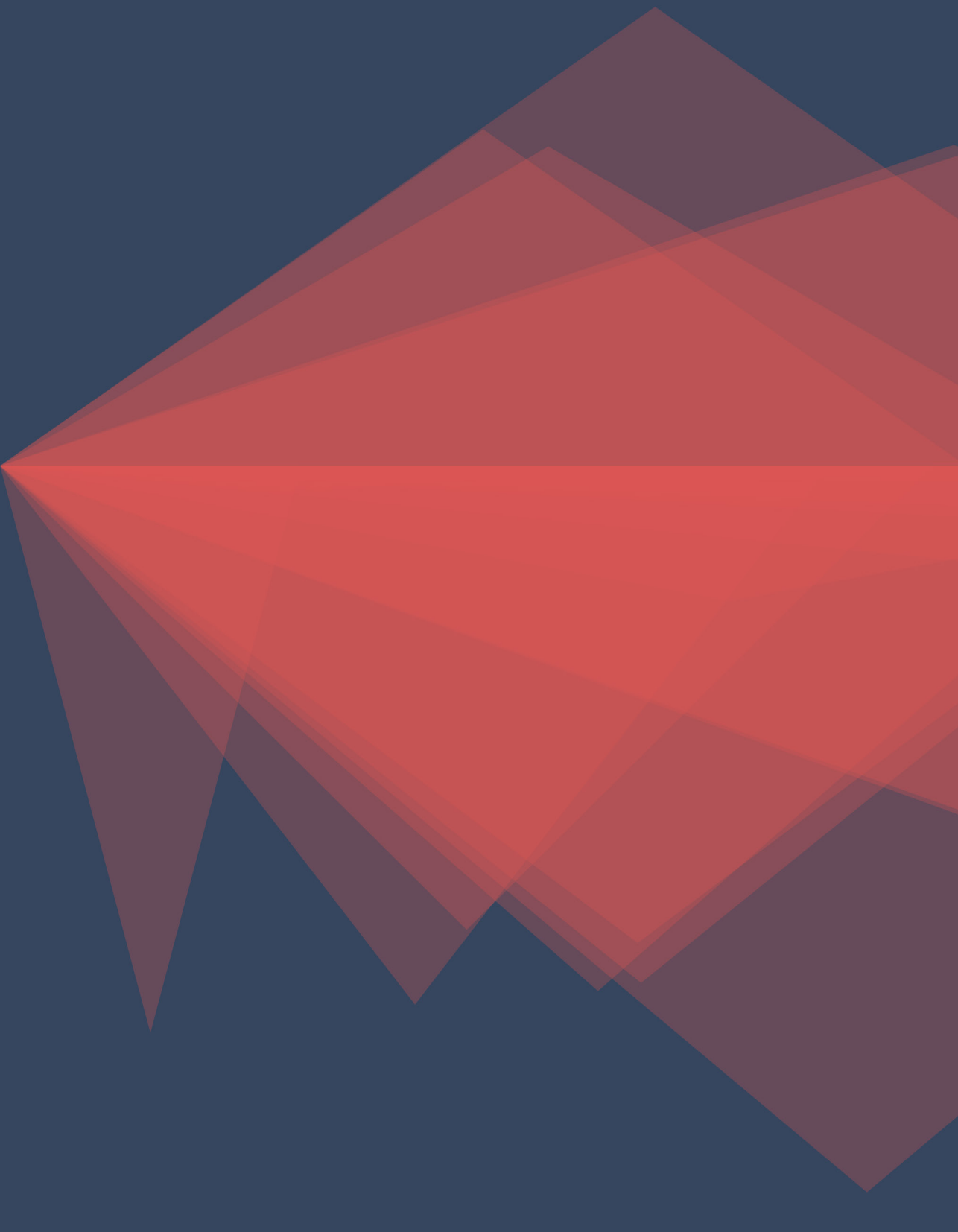
(33) Selman, M. H. J.; Hemayatkar, M.; Deelder, A. M. A. M.; Wuhrer, M. Cotton HILIC SPE Microtips for Microscale Purification and Enrichment of Glycans and Glycopeptides. Anal. Chem. 2011, 83 (7), 2492–2499. https://doi.org/10.1021/ac1027116.

(34) Bladergroen, M. R.; Reiding, K. R.; Hipgrave Ederveen, A. L.; Vreeker, G. C. M.; Clerc, F.; Holst, S.; Bondt, A.; Wuhrer, M.; van der Burgt, Y. E. M. Automation of High-Throughput Mass Spectrometry-Based Plasma N -Glycome Analysis with Linkage-Specific Sialic Acid Esterification. J. Proteome Res. 2015, 14 (9), 4080–4086. https://doi.org/10.1021/acs.jproteome.5b00538.

(35) Deelder, A. M.; Miller, R. L.; De Jonge, N.; Krijger, F. W. Detection of Schistosome Antigen in Mummies. Lancet 1990, 335 (8691), 724–725. https://doi.org/10.1016/0140-6736(90)90838-V.

(36) Anderson, L. Six Decades Searching for Meaning in the Proteome. J. Proteomics 2014, 107, 24–30. https://doi.org/10.1016/j.jprot.2014.03.005.

(37) Anderson, L. Within Sight of a Rational Pipeline for Development of Protein Diagnostics. Clin. Chem. 2012, 58 (1), 28–30. https://doi.org/10.1373/clinchem.2011.173377.



CHAPTER 4

O- and *N*-glycosylation analysis of cell lines by ultrahigh resolution MALDI-FTICR-MS

Based on:

Vreeker, G.C.M.[1,2], *Nicolardi, S.*[1,2], *Madunic, K.*[1], *Kotsias, M.*[1,3], *Van der Burgt, Y.E.M.*[1], *Wuhrer, M.*[1]*: O- and N-glycosylation analysis from of cell lines by ultrahigh resolution MALDI-FTICR-MS. Int. J. Mass Spectrom. 448, 116267 (2020). doi:10.1016/j.ijms.2019.116267. Supplemental info available.*

1 Center for Proteomics and Metabolomics, Leiden University Medical Center, The Netherlands

2 Department of Surgery, Leiden University Medical Center, The Netherlands

3 Ludger Ltd, Culham Science Centre, Abingdon, Oxfordshire, United Kingdom

Abstract

Glycosylation analysis from biological samples is often challenging due to the high complexity of the glycan structures found in these samples. In the present study *N*- and *O*- glycans from human colorectal cancer cell lines and human plasma were analyzed using ultrahigh resolution MALDI-FTICR-MS. *N*-glycans were enzymatically released from cell lines and plasma proteins, whereas beta-elimination was used for the release of *O*-glycans from the cells. The purified samples were mass analyzed using a 15T MALDI-FTICR-MS system, with additional MS/MS (collision-induced dissociation) experiments for *O*-glycan identifications. A total of 104 *O*-glycan and 62 *N*-glycan compositions were observed in the spectra obtained from colorectal cancer cell line samples. In the cell line *N*-glycan spectra, the highest intensity signals originated from high-mannose glycans, next to the presence of various complex type glycans. Notably, in the *O*-glycan spectra mono- and disaccharide signals were observed, which are difficult to detect using alternative glycomic platforms such as porous graphitized carbon LC-MS. In the *N*-glycan spectra from plasma, isobaric species were resolved in MALDI-FTICR-MS spectra using absorption mode whereas these overlapped in magnitude mode. The use of ultrahigh resolution MALDI-FTICR-MS for the analysis of glycans in complex mixtures enables us to confidently analyze glycans in the matrix region of the spectrum and to differentiate isobaric glycan species.

Introduction

Protein glycosylation is a common co- and post-translational modification that changes protein structure and function, and as a consequence plays an important role in various biological processes.[1–4] With regard to clinical utility and anticipated diagnostic applications, protein glycosylation changes have been studied for many different health conditions, aging, multiple types of cancer and autoimmune diseases.[5–9] The two main types of protein glycosylation are *N*- and mucin-type *O*-glycosylation.

Most *N*-glycans share a pentasaccharide core-structure and are attached to a specific asparagine in the protein backbone (consensus sequence Asn-X-Ser/Thr with X any amino acid except proline). *N*-glycans are often large, complex structures that have been studied extensively relying on their enzymatic release from protein backbones.[10] To study protein *N*-glycosylation various analytical methods are commonly used, including fluorescent labeling and detection, capillary electrophoresis, liquid chromatography (LC) and mass spectrometry (MS).[6,11–21]

Mucin-type *O*-glycans often vary vastly in size, from GalNAc monosaccharides attached to serine or threonine residues (Tn-antigens) to large oligosaccharides exhibiting a vast range of terminal glycan motifs. Notably, for a comprehensive analysis of *O*-glycans, a broad specificity release enzyme is lacking [20,22] and the analysis of *O*-glycans often relies on their chemical release from the protein backbone by reductive beta-elimination, resulting in glycan alditols that lack the reducing end for labeling with e.g. a fluorescent tag.[23] This makes MS-analysis the preferred method to characterize *O*-glycans, for example in combination with porous graphitized carbon (PGC)-LC-MS or as permethylated glycans using matrix-assisted laser desorption/ionization (MALDI)-MS.[12,24] It is noted that in both the PGC LC-MS workflows and in MALDI-MS the detection of short-chain *O*-glycans (mono- and disaccharide units) is prohibited because of loss during sample preparation and the overlap with matrix peaks, respectively. Of note, these small glycans are functionally relevant in cancer[25,26], and consequently methods are needed that allow the facile detection and characterization of these *O*-glycans from biological samples.

For glycan characterizations MALDI-MS is commonly performed on a time-of-flight (TOF) mass analyzer. However, the limited resolving power hampers identification of isobaric species in MALDI-TOF spectra.[27] In the case of *N*-glycans peak overlapping has been observed for large tri- and tetraantennary glycans[27] while for *O*-glycans high resolving power (RP) is needed to distinguish small structures from MALDI matrix signals. The highest RP and mass accuracy is obtained with Fourier transform ion cyclotron resonance mass spectrometer (FTICR-MS).[28] The benefits of ultrahigh RPs have been demonstrated for the analysis of complex mixtures.[29,30]

However, such high RPs generally come at the price of decreased sensitivities and longer measurement times.[27] Therefore, increasing RP post-acquisition by means of data processing is an attractive strategy.

The objective of the present study is to apply ultrahigh resolution MALDI-FTICR-MS for detailed *N*- and *O*-glycosylation analysis of complex samples. As an example, two cancer cell lines are selected, from which *N*-glycosylation MALDI-TOF profiles were previously reported by our group. In the current study the high RP of MALDI-FTICR-MS obtained in absorption-mode provides glycan profiles with low ppm mass precision that facilitates spectral alignment of multiple measurements or different samples. High RP of MALDI-FTICR-MS obtained in absorption mode allows the resolution of isobaric glycan species and provides glycan profiles with low mass measurement errors that facilitates spectral alignment of multiple measurements or different samples.[31] Furthermore, high mass accuracy over a large *m*/*z*-range improves *N*-glycan identifications. In addition, this study focuses on small *O*-glycans (mono- and disaccharides) in the matrix region of the MALDI-FTICR-MS spectra, to provide a complimentary strategy to porous graphitized carbon-LC-MS analysis.

Materials and Methods

Materials

Plasma standard (Visucon-F frozen normal control plasma, pooled from 20 human donors, citrated and buffered with 0.02 M HEPES) was obtained from Affinity Biologicals (Ancaster, ON, Canada). Cancer cell lines SW48 and SW1116 were obtained from the Department of Surgery at Leiden University Medical Center (Leiden, The Netherlands). The 1.5 mL Eppendorf® Safe-Lock microcentrifuge tubes (Eppendorf tubes), potassium hydroxide (KOH), potassium borohydride (KBH_4), glacial acetic acid, methanol (MeOH), mucin from bovine submaxillary glands (BSM) type I-S and Iodomethane (ICH_3) were obtained from Sigma (Dorset, UK). The 96-well release plates (4ti-0125), the PCR plates, the foil pierce seals, the semi-automatic heat sealer (HT121TS), the polypropylene collection plates and the silicone plate lids were purchased from 4titude (Surrey, UK). The LudgerTagTM permethylation microplate kit (LT-PERMET-VP96) was obtained from Ludger Ltd (Oxfordshire, UK). The VersaPlate tubes were purchased from Agilent Technologies (Stockport, UK). The CEX resin (BIO-RAD, AG® 50W-X2 Resin) was purchased from Charlton Scientific (Oxon, UK). Trifluoroacetic acid (TFA), fetuin from fetal calf serum, tris(hydroxymethyl)amino-methane, sodium borohydride ($NaBH_4$), Hydrochloric acid (HCl), DL-Dithiothreitol (DTT), 8 M guanidine hydrochloride (GuHCl), ammonium bicarbonate, 50% sodium hydroxide (NaOH), 1-hydroxybenzotriazole 97% (HOBt), super-DHB (9:1 mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid, sDHB), cation exchange resin Dowex 50W X8; cat. no. 217514 and ammonium acetate were obtained from Sigma Aldrich (Steinheim, Germany) and analytical grade ethanol, sodium chloride (NaCl) and methanol were purchased from Merck (Darmstadt, Germany). 1-Ethyl-3-(3-(dimethylamino)-propyl) carbodiimide (EDC) hydrochloride was purchased at Fluorochem (Hadfield, UK) and HPLC-grade acetonitrile (ACN) was obtained from Biosolve (Valkenswaard, The Netherlands). Ethylenediaminetetraacetic acid solution pH 8.0 and potassium hydroxide (KOH) were obtained from Fluka and PNGase F (Flavobacterium meningosepticum recombinant in E. coli; cat. no. 11365193001) and complete EDTA free protease inhibitor cocktail tablets were purchased at Roche. MultiScreen® HTS 96-multiwell plates (pore size 0.45 μm) with high protein-binding membrane (hydrophobic Immobilon-P PVDF membrane) were purchased from Millipore, 96-well PP filter plate, cat.no. OF1100 from Orochem technologies (Naperville, USA), α2,3-neuraminidase, cat. no. GK80040) from New England Biolabs (Ipswich, MA, USA), SPE bulk sorbent Carbograph from Grace discovery sciences (Columbia, USA), Hepes-buffered RPMI 1640 culture media was purchased from Gibco (Paisley, UK). 0.5% trypsin-EDTA solution 10x was obtained from Santa Cruz Biotechnology

(Dallas, USA), fetal bovine serum (FBS) and penicillin/streptomycin at Invitrogen (Carlsbad, USA). T75 cell culture flasks were purchased at Greiner-Bio One B.V. (Alphen aan de Rijn, The Netherlands). Milli-Q water (MQ) was generated from a QGard 2 system (Millipore, Amsterdam, The Netherlands), which was maintained at ≥18 MΩ. All automated steps in the analytical workflow described were performed using a Hamilton MICROLAB STARlet Liquid Handling Workstation from Hamilton Robotics Inc. (Bonaduz, Switzerland).

Cell culture

The cells were cultured as described in Holst et al.[32] Pellets from three technical replicates of each cancer cell (~2 x 1,000,000 cells/technical replicate) were transferred into 1.5 mL Eppendorf tubes, and stored at -20 °C.

***O*-glycan release and purification from cell lines**

Cell pellets containing ca. 2 million cells were resuspended by pipetting action in 100 µL of water with resistivity 18.2 MΩ and homogenized for 90 min in a sonication bath. Following the disruption of cellular plasma membranes, each sample was manually transferred into a reaction well on the 96-well release plate. *O*-glycans were chemically cleaved from cell line proteins by semi-automated reductive β-elimination.[33] Briefly, the 96-well release plate containing the samples in-solution was placed on the robot deck. 40 µL of a 1 M KBH_4 solution in 0.1 M KOH in water was dispensed into each reaction well and the contents were mixed by pipetting action. The release plate was sealed with a foil pierce seal in a semi-automatic heat sealer and incubated in an ultrasonic bath at 60 °C for 4 hours off the robot deck. Following the incubation step, the release plate was briefly centrifuged in order to avoid cross contamination of samples. The seal was removed and the release plate was placed back on the robot deck where two aliquots of respectively 2 µL and 30 µL of glacial acetic acid were added to each reaction well to terminate the reductive β-elimination reaction.

Released *O*-glycans were purified by automated cation-exchange (CEX) cleanup, collected in a polypropylene collection plate and dried down in a centrifugal evaporator.[33] Following the automated cation-exchange cleanup, the polypropylene collection plate containing the dried samples was placed back onto the robot deck. A 1 mL aliquot of MeOH was dispensed into each well and the contents were mixed by pipetting action. Samples were dried down in a centrifugal evaporator.

Automated HT permethylation, liquid-liquid extraction (LLE) and MALDI-spotting

The released and purified *O*-glycans were permethylated using the LudgerTagTM

permethylation microplate kit (LT-PERMET-VP96), purified by liquid-liquid extraction, dried down in a centrifugal evaporator and transferred to a PCR plate as previously described.[20] Permethylation and sample purification by liquid-liquid extraction were performed twice on the Hamilton MICROLAB STARlet Liquid Handling Workstation in order to reduce the percentage of partial permethylation and to improve the efficiency of the automated HT permethylation.[20] Samples were again dried down in a centrifugal evaporator in order to be concentrated prior to MALDI-FTICR-MS analysis. The samples were re-suspended in 10 μL of 70% methanol. The sDHB matrix (1.0 μL of 5 mg/mL super DHB with 1 mM NaOH in 50 % ACN) was spotted on a MTP Anchor-Chip 800/384 MALDI-target (Bruker Daltonics) plate followed by 0.5 μL of permethylated sample and allowed to dry. Data acquisition was performed once for each sample on the MALDI-target plate using a 15T SolariX MALDI-FTICR-MS (Bruker Daltonics).

_N_-glycan release from cell lines

N-glycans were released from cell lysates of three technical replicates per cell line using a PVDF- membrane based approach as described previously.[32] Briefly, cell pellets were resuspended in lysis buffer (containing 50 mM Tris HCl, 100 mM NaCl, 1 mM EDTA, Protease inhibitor cocktail), sonicated for 1 h at 60 °C and proteins from ca. 500 000 cells were immobilized on a 96- well PVDF membrane filter plates in denaturing conditions by adding 72,5 μl of 8M GuHCl and 2,5 μl of 200 mM DTT and incubating for 1 hour at 60 °C. After washing the samples with water to remove all the unbound material, *N*-glycans were released by adding 2 μL of PNGase F diluted in 13 μL of MQ to each well, incubating for 5 min, adding another 15 μL of MQ and incubating overnight at 37 °C. The glycans were recovered by centrifugation and washing (3x 30 μL MQ) of the PVDF membrane. Subsequently 20 μL of ammonium acetate (pH 5) was added and incubated for 1h. *N*-glycan reduction was performed as described before.[12] In short, 30 μL of 1 M $NaBH_4$ in 50 mM KOH was added to the released glycans, incubated for 3 h, where after the reaction was quenched with 2 μl of glacial acetic acid. The glycans were desalted using strongly acidic cation exchange resin. 100 μL of resin in methanol was packed in the 96-well PP filter plate and three times washed with each of the following solutions: 100 μl of 1M HCl, 100 μl methanol and 100 uL water. The samples were loaded and elution was performed twice with 40 μL of MQ and centrifugation. Subsequently the samples were dried followed by borate removal with repeated methanol additions and drying (co-evaporation). Additional carbon solid phase extraction step was performed as described previously.[12]

_N_-glycan derivatization, purification and MALDI-spotting

The dried samples were dissolved in 25 μL of MQ water. Then, 5 μL of this sample was added to 50 μL of ethylation reagent (0.25 M HOBt and 0.25 M EDC in EtOH) and incubated for one hour at 37 °C. After incubation 55 μL of ACN was added. Purification was performed as described before.[27,34,35] In short, in-house developed cotton-HILIC microtips were washed three times with 20 μL MQ and subsequently conditioned with three times 20 μL of 85% ACN. The sample was pipetted up-and-down twenty times in the cotton-HILIC tip and thereafter washed three times with 20 μL 85% ACN with 1% TFA and three times with 20 μL 85% ACN. To elute, 10 μL of MQ water was pipetted up-and-down five times. One microliter of sDHB-matrix (5 mg/mL with 1 mM NaOH in 50% ACN) was spotted onto a MALDI target plate (800/384 MTP AnchorChip, Bruker Daltonics, Bremen, Germany) with 5 μL of sample. The spots were left to dry and afterwards measured on a 15T SolariX MALDI-FTICR-MS instrument.

_N_-glycan analysis from plasma

Samples were prepared according to the previously published method.[27] Briefly, the *N*-glycans were released from 6 μL Visucon plasma by first incubating the plasma for 10 min at 60 °C in 12 μL 2% SDS, where after the release mixture (6 μL acidified PBS, 6 μL 4% NP-40 and 0.6 μL PNGase F) was added and the sample was incubated overnight at 37 °C.

Automated derivatization, purification and MALDI-target plate spotting were performed using the liquid handling platform described before.[27,36] The carboxylic acid moieties of the sialic acids were derivatized by adding 2 μL of sample to 40 μL ethylation reagent (0.25 M EDC with 0.25 M HOBt in ethanol) and incubating for one hour at 37 °C. Subsequently 42 μL 100% ACN was added and after 10 minutes purification was performed in the same manner as described in above, however, all amounts were doubled in the automated purification. MALDI-target spotting was performed as described in above, but the sample and matrix volumes spotted were 2 μL and 1 μL respectively with premixing before spotting.

MALDI FT-ICR MS and MS/MS measurements

All MALDI-FTICR-MS measurements were performed on a 15T SolariX XR mass spectrometer (Bruker Daltonics) equipped with a CombiSource, a ParaCell and a SmartBeam-II laser system.

N-glycans were measured as previously reported. Briefly, for each MALDI spot, one spectrum consisted of ten acquired scans and 1 M data points was generated in the *m*/*z*-range 1011.86 to 5000.00. The laser was operated at a frequency of 500 Hz using the “medium” predefined shot pattern and 200 laser shots were

collected per raster (i.e. 2000 per MALDI spot). The ParaCell parameters were as follows: both shimming and gated injection DC bias at 0°, 90°, 180°, and 270° were 9.20, 9.13, 9.20, 9.27 V, respectively; the trapping potentials were set at 9.50 and 9.45 V and the excitation power and sweep step time at 55% and 15μs. The transfer time of the ICR cell was 1.0 ms, and the quadrupole mass filter was set at *m*/*z* 850.

O-glycans spectra were generated from ten acquired scans and 1 M data points, in the *m*/*z*-range 207-5000. As for *N*-glycans, 200 laser shots were collected per raster using a 500 Hz laser frequency and the "medium" predefined shot pattern. The ParaCell parameters were as follows: both shimming and gated injection DC bias at 0°, 90°, 180°, and 270° were 1.20, 1.13, 1.20, 1.27 V, respectively; the trapping potentials were set at 1.50 and 1.45 V and the excitation power and sweep step time at 45% and 15μs. The transfer time of the ICR cell was 0.85 ms, and the quadrupole mass filter was set at *m*/*z* 100. Collision-induced dissociation (CID) experiments were performed with quadrupole (Q1) mass, isolation window and collision energy optimized for each precursor ion.

Absorption mode MALDI-FTICR-MS spectra were obtained using AutoVectis software (Spectroswiss, Lausanne, Switzerland).[37] DataAnalysis Software 5.0 (Bruker Daltonics) and mMass38 were used for the visualization and data analysis of MALDI-(CID)-FTICR-MS spectra.

Results and Discussion

N- and *O*-glycan analysis with MALDI-FTICR-MS

Ultrahigh resolution MALDI-FTICR mass spectra were obtained from *N*- and *O*-glycans from two cell lines (referred to as SW48 and SW1116). In Figure 1 annotated spectra of these cell lines for both glycosylation types are shown. The highest intensity signals found in the *N*-glycan spectra corresponded to high-mannose glycan structures. In addition, complex type glycans were annotated, although observed at lower intensities. The total number of *N*-glycans detected in the spectra from SW48 and SW1116 were 76 and 104, respectively (see Table S-1). Of all *N*-glycan assignments 94% were within a mass measurement error of ±2 ppm , with the remaining 6% within ±4 ppm, in a mass range of *m*/*z* 1000 to 3500. Remarkably, these high-accuracy assignments were reproducibly obtained from a single mass spectrum with lowest and highest glycan peak intensities varying by more than a factor of 5000.

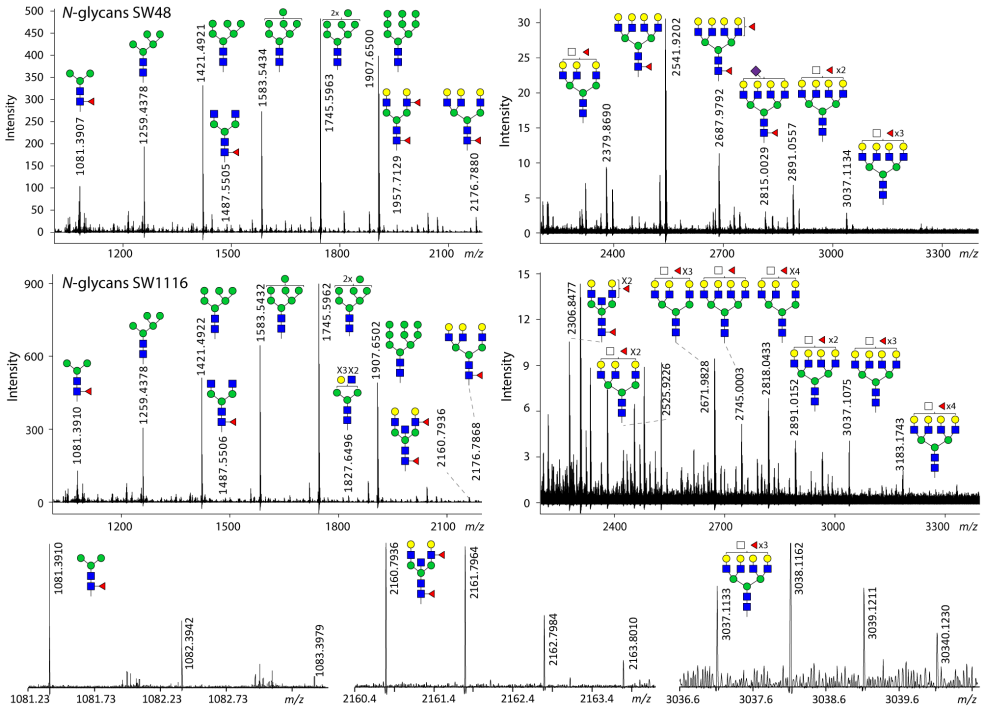


Figure 1. Absorption mode MALDI-FTICR mass spectra of released and reduced *N*-glycans from cell lines SW48 and SW1116. All glycans are assigned as $[M+Na]^+$. Only the relevant *m*/*z*-range for glycans is shown.

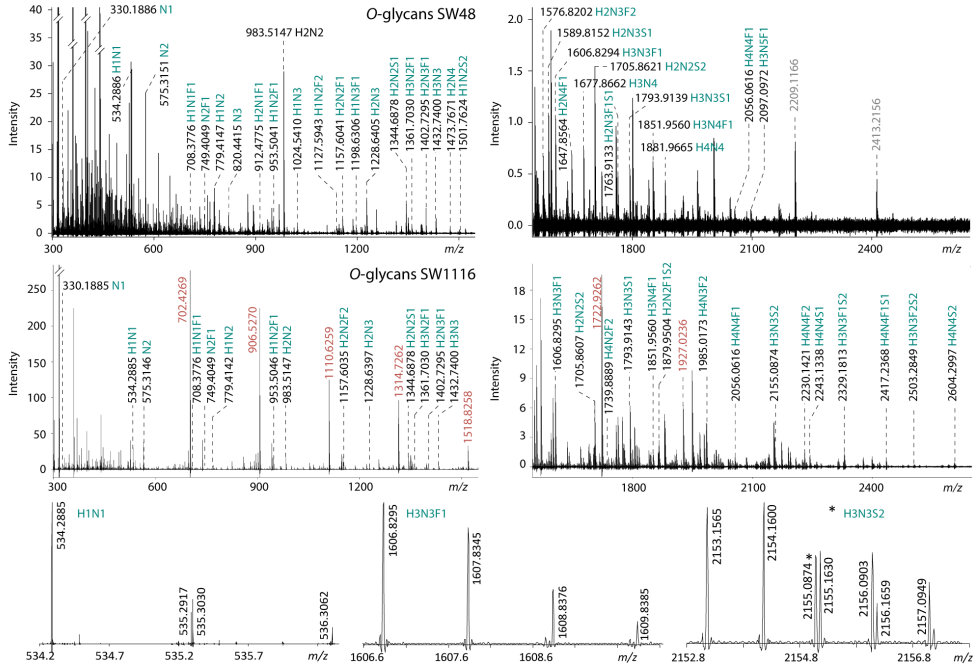


Figure 2. Absorption mode MALDI-FTICR mass spectra of released, reduced and permethylated *O*-glycans from cell lines SW48 and SW1116. All glycan are assigned as $[M+Na]^+$, except for the hexose ladder (in red) in the *O*-glycan spectrum of SW1116, which carries a quaternary trimethyl ammonia group. Only the relevant *m*/*z*-range for glycans is shown.

In the *O*-glycan spectra of SW48 and SW1116 (Figure 2), 44 and 62 different *O*-glycan compositions were annotated respectively (see Table S-2). Of all *O*-glycan assignments 76% were within a mass measurement error of ±2 ppm and an additional 18% within ±6 ppm, in a mass range of *m*/*z* 300 to 3000. Similar to the spectra from the *N*-glycans the peak intensities varied by more than a factor of 5000. For many of these *O*-glycan compositions multiple glycan isomers are possible. This is exemplified in the fragmentation spectra of the sodiated permethylated alditol of the glycan composition H2N2 (Figure 3). Diagnostic fragment ions are observed for both a linear and a branched tetrasaccharide structure (*m*/*z* 690.3291 and *m*/*z* 284.1459, respectively), and it is therefore concluded that this mass spectrometric signal represents a mixture of isomers. The linear glycan is proposed to be a core-1 *O*-glycan structure (Gal1-3/4GlcNAc1-3Gal1-3GalNAcitol) while the branched isomer is presumably a core-2 *O*-glycan structure (Gal1-3/4GlcNAc1-6[Gal1-3] GalNAcitol). In Figure S-1 additional fragmentation spectra are shown that demonstrate similar types of isomeric mixtures. Furthermore, high RPs in MALDI-FTICR spectra allowed detection of glycan signals in the matrix region. An interesting

signal that was found corresponds to a single permethylated *N*-acetylhexosaminitol (*m*/*z* 330.19). This *N*-acetylhexosaminitol is most likely derived from a single GalNAc mucin-type *O*-glycan (Tn-antigen). Importantly, the Tn antigen cannot be detected with current PGC-LC-MS workflows for negative-mode analysis of native (i.e. non-permethylated) *O*-glycan alditols[12], and the mass spectrometric analysis of permethylated *O*-glycan alditols therefore represents a complementary technique making these biologically important short-chain *O*-glycans amenable for analysis. Next to the here presented MALDI-FTICR-MS analysis, the LC-MS analysis of permethylated glycans by reversed phase and porous graphitized carbon stationary phases has recently obtained increased attention. Such platforms may complement the MALDI-FTICR-MS approaches by providing isomer separation, as has recently been described for *N*-glycans.[39–42]

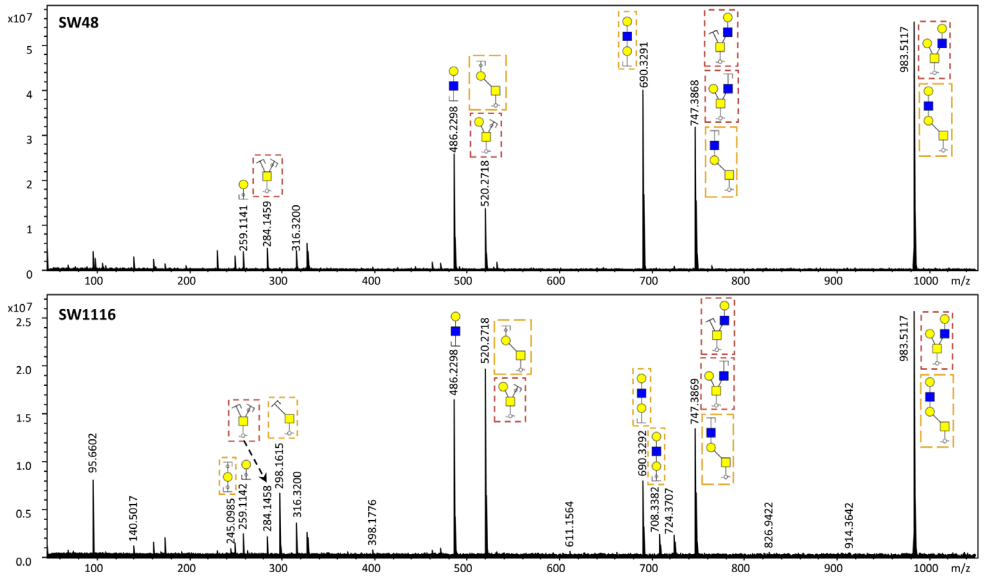


Figure 3. Fragmentation spectrum of *m*/*z* 983.5117 which is assigned with the composition H2N2. The diagnostic ions (e.g. *m*/*z* 245, 284, 298, 690, 708) indicate that there are two different isomers of H2N2 present in the cell line samples.

Notably, it is still unclear to which extent other HexNAc sources such as protein-linked *N*-acetylglucosamine (*O*-GlcNAc) which is known to be mainly present on cytosolic proteins[43], may be contributing to the *N*-acetylhexosaminitol signal. MALDI-FTICR MS/MS analysis of permethylated *N*-acetylglucosaminitol and *N*-aceetylgalactosaminitol standards resulted in a similar fragmentation pattern for both species (data not shown), pointing out that additional experiments are needed in order to identify whether or not this signal is confounded by *O*-GlcNAc,

and whether it can be largely attributed to the Tn-antigen.
Moreover, in Figure S-2 the full mass range of one *O*-glycan spectrum is shown, with a highly intense peak at *m*/*z* 316.32 that could be identified by MS/MS as dodecyl-bis(2-hydroxyethyl)-methylazanium, a permanently charged compound originating from plastics. Further optimizations in the sample workup and purifications are therefore needed to turn these *O*-glycan spectra into robust profiles.

As a final remark on the *O*-glycan spectra, it was noted that the most intense peaks in SW1116 corresponded to a hexose ladder (see Figure 2). These signals were also observed in SW48 at a lower intensity. The hexose ladder can be explained by the release of part of the *N*-glycans in the beta-elimination reaction, resulting in the appearance of additional high mannose glycans, as these *N*-glycans were identified amongst the most abundant species in this cell line. The ladder contains up to nine hexoses and a remainder mass that was identified as a hexosamine with a quaternary methylated amine group, thus carrying a positive charge (see fragmentation data in Figure S-3). An overview of the observed *m*/*z* values from the hexose ladder is shown in Table S-3. An additional confirmation for the hypothesis that the hexose ladder resulted from the high-mannose glycans was provided by the observation of two additional hexose ladders in the *O*-glycan spectra: both ladders likewise consisted of up to nine hexoses with a *N*-acetylhexosamine or a hexosamine, which both ionized as $[M+Na]^+$. The observed *m*/*z* values of these ladders are also shown in Table S-3.

Resolving glycan signals

The measurements on *N*- and *O*-glycans from cell lines were compared with previous measurements performed in our group on released *N*-glycans from plasma.[27] Moreover, in the current study we evaluated the *N*-glycan MALDI-FTICR data from plasma in absorption mode instead of magnitude mode, resulting in a further improved resolution. Previously we reported that isobaric structures such as H7N6F1L1E1 and H6N5F1E3 were not distinguishable in MALDI-time-of-flight (TOF)-MS, and that overlapping signals were observed with magnitude mode measurements in MALDI-FTICR-MS.[27] Now, with absorption mode evaluation of the signals these two glycans are fully resolved (see Figure 4). In the *O*-glycans spectra from cell lines this same effect is shown (Figure 4 and S-4). The signal, assigned as H3N3F1S1, is in absorption mode separated from the overlapping signal in magnitude mode.

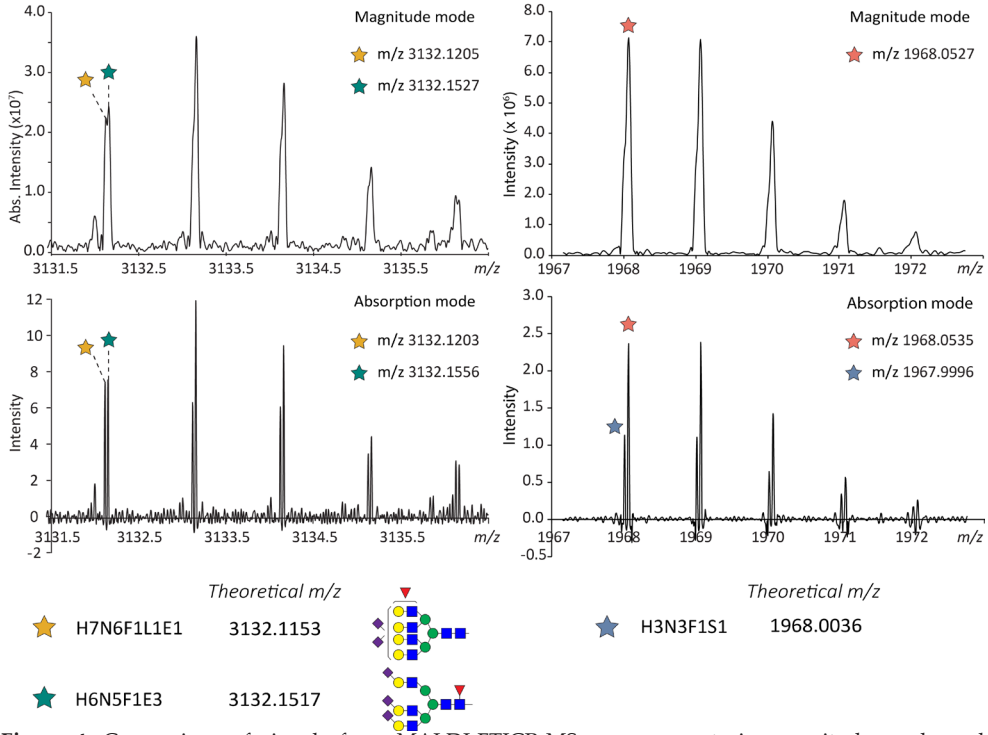


Figure 4. Comparison of signals from MALDI-FTICR-MS measurements in magnitude mode and absorption mode. Due to increased resolution in absorption mode measurements the overlapping peaks are resolved. In de left panel two *N*-glycans are resolved and in the right panel an underlying *O*-glycan signal is resolved from an overlapping non-glycan analyte.

Conclusion

This study demonstrated that MALDI-FTICR-MS analysis allows for glycan identifications in complex mass spectra with low ppm mass measurement errors. However, a confident assignment of monosaccharide composition will not resolve the issue that complex biological samples often contain multiple glycan isomers. To this end diagnostic fragment ions in CID-spectra assist structure elucidation, as was exemplified in this study. In case overlapping signals were observed in MALDI-FTICR-MS absorption mode visualizations were a useful approach for increasing resolution and improving *O*- and *N*-glycan identifications. The benefits of ultrahigh RPs allow the analysis of complex samples, such as a *N*- or *O*-glycome, without the need for time consuming chromatography. The here presented strategy with low ppm mass measurement errors and confident glycan compositional assignment holds great potential for a comparative study on patient sample cohorts, such as plasma samples or cell lines.

Acknowledgements

The society "Genootschap ter ondersteuning van de vroege opsporing van kanker" (Lisse, The Netherlands) financially supported parts of this research performed at the LUMC to further endorse the development of a blood-based test for early detection of cancer (no grant number applicable). Katarina Madunic and Maximilianos Kotsias were supported by the European Commission GlyCoCan project (contract no. 676412).

References

(1) Geyer, H.; Geyer, R. Strategies for Analysis of Glycoprotein Glycosylation. Biochim. Biophys. Acta - Proteins Proteomics 2006, 1764 (12), 1853–1869. https://doi.org/10.1016/j.bbapap.2006.10.007.

(2) Cummings, R. D. The Repertoire of Glycan Determinants in the Human Glycome. Mol. Biosyst. 2009, 5 (10), 1087. https://doi.org/10.1039/b907931a.

(3) Dell, A.; Morris, H. R. Glycoprotein Structure Determination by Mass Spectrometry. Science. March 23, 2001, pp 2351–2356. https://doi.org/10.1126/science.1058890.

(4) Spiro, R. G. Protein Glycosylation: Nature, Distribution, Enzymatic Formation, and Disease Implications of Glycopeptide Bonds. Glycobiology 2002, 12 (4), 43R-56R.

(5) Bondt, A.; Rombouts, Y.; Selman, M. H. J.; Hensbergen, P. J.; Reiding, K. R.; Hazes, J. M. W.; Dolhain, R. J. E. M.; Wuhrer, M. IgG Fab Glycosylation Analysis Using a New Mass Spectrometric High-Throughput Profiling Method Reveals Pregnancy-Associated Changes. Mol. Cell. Proteomics 2014, 31 (0), 1–30. https://doi.org/10.1074/mcp.M114.039537.

(6) Reiding, K. R.; Bondt, A.; Hennig, R.; Gardner, R. A.; O'Flaherty, R.; Trbojević-Akmačić, I.; Shubhakar, A.; Hazes, J. M. W.; Reichl, U.; Fernandes, D. L.; et al. High-Throughput Serum N -Glycomics: Method Comparison and Application to Study Rheumatoid Arthritis and Pregnancy-Associated Changes. Mol. Cell. Proteomics 2019, 18 (1), 3–15. https://doi.org/10.1074/mcp.RA117.000454.

(7) Saldova, R.; Wormald, M. R.; Dwek, R. A.; Rudd, P. M. Glycosylation Changes on Serum Glycoproteins in Ovarian Cancer May Contribute to Disease Pathogenesis. Dis. Markers 2008, 25 (4–5), 219–232. https://doi.org/10.1155/2008/601583.

(8) Alley, W. R.; Madera, M.; Mechref, Y.; Novotny, M. V. Chip-Based Reversed-Phase Liquid Chromatography-Mass Spectrometry of Permethylated N-Linked Glycans: A Potential Methodology for Cancer-Biomarker Discovery. Anal. Chem. 2010, 82 (12), 5095–5106. https://doi.org/10.1021/ac100131e.

(9) Ruhaak, L. R.; Uh, H. W.; Beekman, M.; Hokke, C. H.; Westendorp, R. G. J.; Houwing-Duistermaat, J.; Wuhrer, M.; Deelder, A. M.; Slagboom, P. E. Plasma Protein N-Glycan Profiles Are Associated with Calendar Age, Familial Longevity and Health. J. Proteome Res. 2011, 10 (4), 1667–1674. https://doi.org/10.1021/pr1009959.

(10) Tarentino, A. L.; Gómez, C. M.; Plummer, T. H. Deglycosylation of Asparagine-Linked Glycans by Peptide: N-Glycosidase F. Biochemistry 1985, 24 (1969), 4665–4671. https://doi.org/10.4052/tigg.5.163.

(11) Zhou, S.; Hu, Y.; Desantos-Garcia, J. L.; Mechref, Y. Quantitation of Permethylated N-Glycans through Multiple-Reaction Monitoring (MRM) LC-MS/MS. J. Am. Soc. Mass Spectrom. 2015, 26 (4), 596–603. https://doi.org/10.1007/s13361-014-1054-1.

(12) Jensen, P. H.; Karlsson, N. G.; Kolarich, D.; Packer, N. H. Structural Analysis of N- and O-Glycans Released from Glycoproteins. Nat. Protoc. 2012, 7 (7), 1299–1310. https://doi.org/10.1038/nprot.2012.063.

(13) Abrahams, J. L.; Campbell, M. P.; Packer, N. H. Building a PGC-LC-MS N-Glycan Retention Library and Elution Mapping Resource. Glycoconj. J. 2018, 35 (1), 15–29. https://doi.org/10.1007/s10719-017-9793-4.

(14) Yang, S.; Clark, D.; Liu, Y.; Li, S.; Zhang, H. High-Throughput Analysis of N-Glycans Using AutoTip via Glycoprotein Immobilization. Sci. Rep. 2017, 7 (1), 1–11. https://doi.org/10.1038/s41598-017-10487-8.

(15) Grünwald-Gruber, C.; Thader, A.; Maresch, D.; Dalik, T.; Altmann, F. Determination of True Ratios of Different N-Glycan Structures in Electrospray Ionization Mass Spectrometry. Anal. Bioanal. Chem. 2017, 409 (10), 2519–2530. https://doi.org/10.1007/s00216-017-0235-8.

(16) Szekrényes, Á.; Park, S. A. S.; Cosgrave, E.; Jones, A.; Haxo, T.; Kimzey, M.; Pourkaveh, S.; Szabó, Z.; Sosic, Z.; Feng, P.; et al. Multi-Site N-Glycan Mapping Study 2: UHPLC. Electrophoresis 2018, 39 (7), 998–1005. https://doi.org/10.1002/elps.201700463.

(17) Peng, Y.; Lv, J.; Yang, L.; Wang, D.; Zhang, Y.; Lu, H. A Streamlined Strategy for Rapid and Selective Analysis of Serum N-Glycome. Anal. Chim. Acta 2019, 1050, 80–87. https://doi.org/10.1016/j.aca.2018.11.002.

(18) Le, H. T.; Park, K. H.; Jung, W.; Park, H. S.; Kim, T. W. Combination of Microwave-Assisted Girard Derivatization with Ionic Liquid Matrix for Sensitive MALDI-TOF MS Analysis of Human Serum N-Glycans. J. Anal. Methods Chem. 2018, 2018. https://doi.org/10.1155/2018/7832987.

(19) Keser, T.; Pavić, T.; Lauc, G.; Gornik, O. Comparison of 2-Aminobenzamide, Procainamide and RapiFluor-MS as Derivatizing Agents for High-Throughput HILIC-UPLC-FLR-MS N-Glycan Analysis. Front. Chem. 2018, 6 (July), 1–12. https://doi.org/10.3389/fchem.2018.00324.

(20) Shubhakar, A.; Kozak, R. P.; Reiding, K. R.; Royle, L.; Spencer, D. I. R.; Fernandes, D. L.; Wuhrer, M. Automated High-Throughput Permethylation for Glycosylation Analysis of Biologics Using MALDI-TOF-MS. Anal. Chem.

2016, 88 (17), 8562–8569. https://doi.org/10.1021/acs.analchem.6b01639.

(21) Hu, Y.; Shihab, T.; Zhou, S.; Wooding, K.; Mechref, Y. LC-MS/MS of Permethylated N-Glycans Derived from Model and Human Blood Serum Glycoproteins. Electrophoresis 2016, 37 (11), 1498–1505. https://doi.org/10.1002/elps.201500560.

(22) Kudelka, M. R.; Ju, T.; Heimburg-Molinaro, J.; Cummings, R. D. Simple Sugars to Complex Disease—Mucin-Type O-Glycans in Cancer. In Advances in Cancer Research; Elsevier Inc., 2015; Vol. 126, pp 53–135. https://doi.org/10.1016/bs.acr.2014.11.002.

(23) Zauner, G.; Kozak, R. P.; Gardner, R. A.; Fernandes, D. L.; Deelder, A. M.; Wuhrer, M. Protein O-Glycosylation Analysis. Biol. Chem. 2012, 393 (8), 687–708. https://doi.org/10.1515/hsz-2012-0144.

(24) Goetz, J. A.; Novotny, M. V.; Mechref, Y. Enzymatic/Chemical Release of O-Glycans Allowing MS Analysis at High Sensitivity. Anal. Chem. 2009, 81 (23), 9546–9552. https://doi.org/10.1021/ac901363h.

(25) Liu, Z.; Liu, J.; Dong, X.; Hu, X.; Jiang, Y.; Li, L.; Du, T.; Yang, L.; Wen, T.; An, G.; et al. Tn Antigen Promotes Human Colorectal Cancer Metastasis via H-Ras Mediated Epithelial-Mesenchymal Transition Activation. J. Cell. Mol. Med. 2019, 23 (3), 2083–2092. https://doi.org/10.1111/jcmm.14117.

(26) Fu, C.; Zhao, H.; Wang, Y.; Cai, H.; Xiao, Y.; Zeng, Y.; Chen, H. Tumor-Associated Antigens: Tn Antigen, STn Antigen, and T Antigen. Hla 2016, 88 (6), 275–286. https://doi.org/10.1111/tan.12900.

(27) Vreeker, G. C. M.; Nicolardi, S.; Bladergroen, M. R.; van der Plas, C. J.; Mesker, W. E.; Tollenaar, R. A. E. M.; van der Burgt, Y. E. M.; Wuhrer, M. Automated Plasma Glycomics with Linkage-Specific Sialic Acid Esterification and Ultrahigh Resolution MS. Anal. Chem. 2018, 90 (20), 11955–11961. https://doi.org/10.1021/acs.analchem.8b02391.

(28) Kilgour, D. P. A.; Wills, R.; Qi, Y.; O'Connor, P. B. Autophaser: An Algorithm for Automated Generation of Absorption Mode Spectra for FT-ICR MS. Anal. Chem. 2013, 85 (8), 3903–3911. https://doi.org/10.1021/ac303289c.

(29) Nicolardi, S.; Bogdanov, B.; Deelder, A. M.; Palmblad, M.; Van Der Burgt, Y. E. M. Developments in FTICR-MS and Its Potential for Body Fluid Signatures. Int. J. Mol. Sci. 2015, 16 (11), 27133–27144. https://doi.org/10.3390/ijms161126012.

(30) Marshall, A. G.; Rodgers, R. P. Petroleomics: Chemistry of the Underworld. Proc. Natl. Acad. Sci. 2008, 105 (47), 18090–18095. https://doi.org/10.1073/pnas.0805069105.

(31) Nicolardi, S.; Palmblad, M.; Hensbergen, P. J.; Tollenaar, R. A. E. M.; Deelder, A. M.; van der Burgt, Y. E. M. Precision Profiling and Identification of Human Serum Peptides Using Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Rapid Commun. mass Spectrom. 2011, 25 (23), 3457–3463. https://doi.org/10.1002/rcm.5246.

(32) Holst, S.; Deuss, A. J. M.; van Pelt, G. W.; van Vliet, S. J.; Garcia-Vallejo, J. J.; Koeleman, C. A. M.; Deelder, A. M.; Mesker, W. E.; Tollenaar, R. A.; Rombouts, Y.; et al. N-Glycosylation Profiling of Colorectal Cancer Cell Lines Reveals Association of Fucosylation with Differentiation and Caudal Type Homebox 1 (CDX1)/Villin MRNA Expression. Mol. Cell. Proteomics 2016, 15 (1), 124–140. https://doi.org/10.1074/mcp.M115.051235.

(33) Kotsias, M.; Kozak, R. P.; Gardner, R. A.; Wuhrer, M.; Spencer, D. I. R. Improved and Semi-Automated Reductive β-Elimination Workflow for Higher Throughput Protein O-Glycosylation Analysis. PLoS One 2019, 14 (1), 1–14. https://doi.org/10.1371/journal.pone.0210759.

(34) Selman, M. H. J.; Hemayatkar, M.; Deelder, A. M. A. M.; Wuhrer, M. Cotton HILIC SPE Microtips for Microscale Purification and Enrichment of Glycans and Glycopeptides. Anal. Chem. 2011, 83 (7), 2492–2499. https://doi.org/10.1021/ac1027116.

(35) Reiding, K. R.; Blank, D.; Kuijper, D. M.; Deelder, A. M.; Wuhrer, M. High-Throughput Profiling of Protein N-Glycosylation by MALDI-TOF-MS Employing Linkage-Specific Sialic Acid Esterification. Anal. Chem. 2014, 86 (12), 5784–5793. https://doi.org/10.1021/ac500335t.

(36) Bladergroen, M. R.; Reiding, K. R.; Hipgrave Ederveen, A. L.; Vreeker, G. C. M.; Clerc, F.; Holst, S.; Bondt, A.; Wuhrer, M.; van der Burgt, Y. E. M. Automation of High-Throughput Mass Spectrometry-Based Plasma N -Glycome Analysis with Linkage-Specific Sialic Acid Esterification. J. Proteome Res. 2015, 14 (9), 4080–4086. https://doi.org/10.1021/acs.jproteome.5b00538.

(37) van der Burgt, Y. E. M.; Kilgour, D. P. A.; Tsybin, Y. O.; Srzentić, K.; Fornelli, L.; Beck, A.; Wuhrer, M.; Nicolardi, S. Structural Analysis of Monoclonal Antibodies by Ultrahigh Resolution MALDI In-Source Decay FT-ICR Mass Spectrometry. Anal. Chem. 2019, 91 (3), 2079–2085. https://doi.org/10.1021/acs.analchem.8b04515.

(38) Strohalm, M.; Hassman, M.; Košata, B.; Kodíček, M. MMass Data Miner: An Open Source Alternative for Mass Spectrometric Data Analysis. Rapid Commun. Mass Spectrom. 2008, 22 (6), 905–908. https://doi.org/10.1002/rcm.3444.

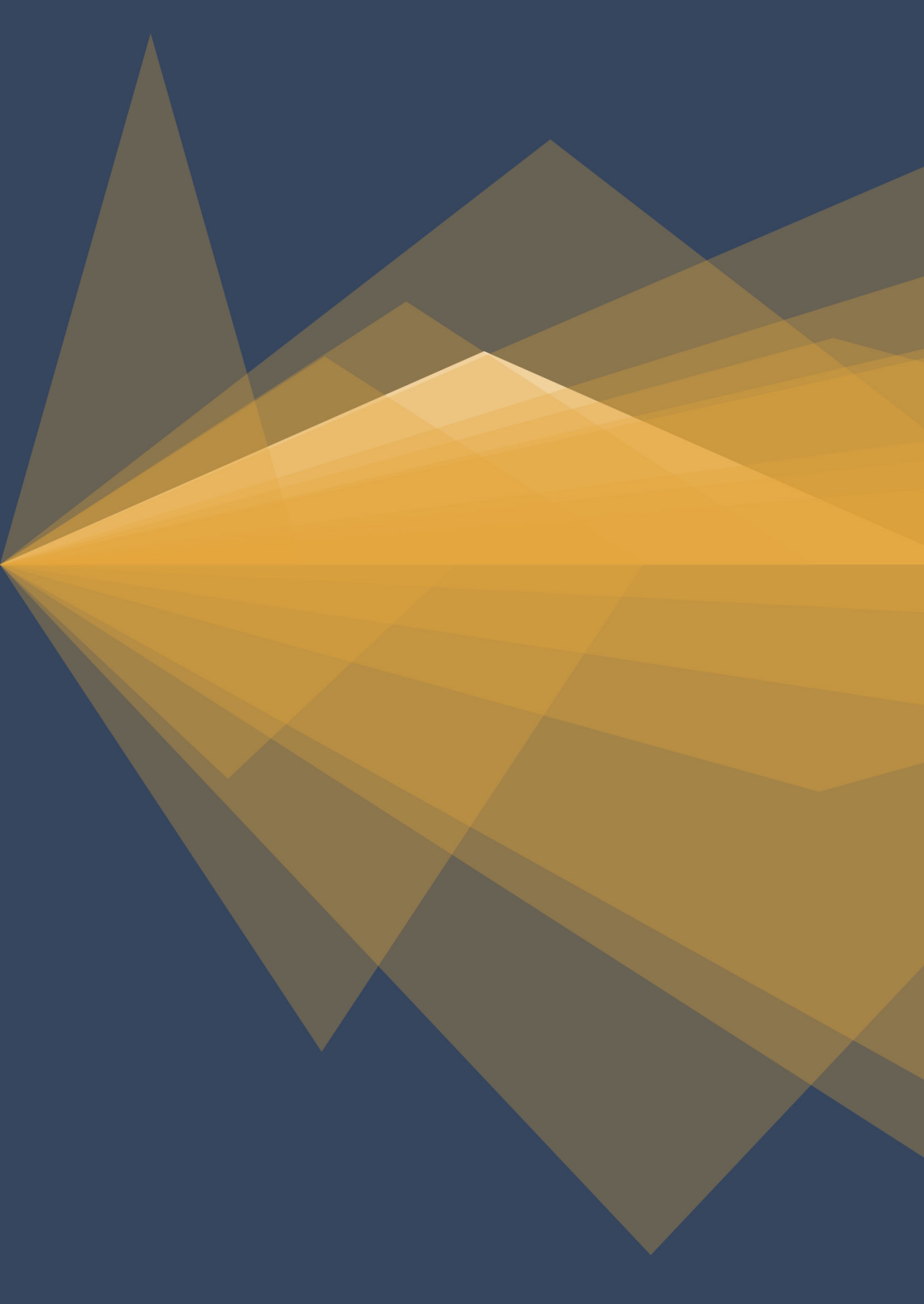
(39) Zhou, S.; Hu, Y.; Mechref, Y. High-Temperature LC-MS/MS of Permethylated Glycans Derived from Glycoproteins. Electrophoresis 2016, 37 (11), 1506–1513. https://doi.org/10.1002/elps.201500568.

(40) Huang, Y.; Zhou, S.; Zhu, J.; Lubman, D. M.; Mechref, Y. LC-MS/MS Isomeric Profiling of Permethylated N-Glycans Derived from Serum Haptoglobin of Hepatocellular Carcinoma (HCC) and Cirrhotic Patients. Electrophoresis 2017, 38 (17), 2160–2167. https://doi.org/10.1002/elps.201700025.

(41) Zhou, S.; Dong, X.; Veillon, L.; Huang, Y.; Mechref, Y. LC-MS/MS Analysis of Permethylated N-Glycans Facilitating Isomeric Characterization. Anal. Bioanal. Chem. 2017, 409 (2), 453–466. https://doi.org/10.1007/s00216-016-9996-8.

(42) Zhou, S.; Wooding, K. M.; Mechref, Y. Analysis of Permethylated Glycan by Liquid Chromatography (LC) and Mass Spectrometry (MS). In High-Throughput Glycomics and Glycoproteomics; Lauc, G., Wuhrer, M., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2017; Vol. 1503, pp 83–96. https://doi.org/10.1007/978-1-4939-6493-2.

(43) Comer, F. I.; Hart, G. W. O -Glycosylation of Nuclear and Cytosolic Proteins . J. Biol. Chem. 2002, 275 (38), 29179–29182. https://doi.org/10.1074/jbc.r000010200.



CHAPTER 5

Serum *N*-Glycome Analysis Reveals Pancreatic Cancer Disease Signatures

Based on:

Vreeker, G.C.M.[1,2]*, *Hanna-Sawires, R.G.*[1]*, *Mohammed, Y.*[2], *Bladergroen, M.R.*[2], *Nicolardi, S.*[1,2], *Dotz, V.*[2], *Nouta, J.*[2], *Bonsing, B.A.*[1], *Mesker, W.E.*[1], *Van der Burgt, Y.E.M.*[2], *Wuhrer, M.*[2]**, *Tollenaar, R.A.E.M.*[1]**: *Serum N-Glycome Analysis Reveals Pancreatic Cancer Disease Signatures. Cancer med. (2020). doi:10.1002/cam4.3439. Supplemental info available.*

1 Department of Surgery, Leiden University Medical Center, Leiden, the Netherlands

2 Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, the Netherlands

*Both authors share first authorship / **Both authors share last authorship

Abstract

Background & Aims: Pancreatic ductal adenocarcinoma (PDAC) is an aggressive cancer type with loco-regional spread that makes the tumor surgically unresectable. Novel diagnostic tools are needed to improve detection of PDAC and increase patient survival. In this study we explore serum protein *N*-glycan profiles from PDAC patients with regard to their applicability to serve as a disease biomarker panel.

Methods: Total serum *N*-glycome analysis was applied to a discovery set (86 PDAC cases/84 controls) followed by independent validation (26 cases/26 controls) using in-house collected serum specimens. Protein *N*-glycan profiles were obtained using ultrahigh resolution mass spectrometry and included linkage-specific sialic acid information. *N*-glycans were relatively quantified and case-control classification performance was evaluated based on glycosylation traits such as branching, fucosylation, and sialylation.

Results: In PDAC patients a higher levelof branching (OR 6.19, p-value 9.21×10-11) and (antenna)fucosylation (OR 13.27, p-value 2.31×10-9) of *N*-glycans was found. Furthermore, the ratio of α2,6- versus α2,3-linked sialylation was higher in patients compared to healthy controls. A classification model built with three glycosylation traits was used for discovery (AUC 0.88) and independent validation (AUC 0.81), with sensitivity and specificity values of 0.85 and 0.71 for the discovery set and 0.75 and 0.72 for the validation set.

Conclusion: Serum *N*-glycome analysis revealed glycosylation differences that allow classification of PDAC patients from healthy controls. It was demonstrated that glycosylation traits rather than single *N*-glycan structures obtained in this clinical glycomics study can serve as a basis for further development of a blood-based diagnostic test.

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is the most common type of pancreatic cancer withan increasing incidence in western countries.[1] Diagnosis of PDAC implies an unfavorable prognosis with five-year survival as low as 5-8%, since the disease is characterized by aggressive local and early metastatic spread. Upon initial diagnosis more than 80% of the tumors is at an advanced stage that does not allow curative resection.[2] Intensive treatment schedules with chemotherapy and/or surgery are associated with complications, side effects and impaired quality of life, while overall survival remains poor.[3] A recent study reported that 10% of the PDAC patients carry a BRCA gene mutation which could provide an opportunity to apply screening and use targeted treatment to improve outcome.[4] It is furthermore noted that the number of PDAC deaths is not far from that of, for example, breast cancer, and a future screening for PDAC may be warranted.[5–9] However, screening programs based on current detection methods that comprise imaging techniques and/or fine-needle-aspirations, are not feasible.[10,11] Moreover, chronic and autoimmune pancreatitis (CP) can mimic PDAC and consequently cause a 5-10% misclassification.[12] For these reasons, new blood-based biomarker tests are pursued that offer a more cost-effective way to detect the disease.[13] This urgent need for additional biomarkers to facilitate clinical decision-making is widely acknowledged, since the only marker available is carbohydrate antigen (CA) 19-9, which is primarily used for patient follow-up (recurrent disease) and has limited value for the detection of PDAC.[14]

Mass spectrometry (MS)-based biomarker studies have shown that post-translational modifications (PTMs) hold potential as an "add-on" to the protein marker or as biomarkers themselves.[13,15,16] It is well-known that a single gene does not transcribe and translate into a single protein but rather in a plethora of proteoforms and that proteome characterization should include the analysis of PTMs.[17] In this context, the relevancy of protein glycosylation has been demonstrated in auto-immune diseases and cancer.[18–21] In-depth glycobiology studies have furthermore revealed the importance of protein glycosylation with regard to folding, trafficking, cell adhesion, recognition processes, and immune response.[22,23] Notably, the previously mentioned marker CA19-9 is a glycan marker, based on a sialyl-Lewis A (sLeA) epitope, which triggered interest in protein glycosylation related to pancreatic cancer. A few studies on *N*-glycosylation profiles in pancreatic cancer have exemplified a biomarker potential, although sample sets were limited.[24–30] Here, we use an automated protocol for the analysis of the total serum *N*-glycome with sialic acid linkage differentiation and high resolution MS[31] and aim for a PDAC disease signature in a discovery cohort with independent validation.

Materials and methods

Patients

Blood samples in the discovery cohort were obtained from 88 patients diagnosed with PDAC and collected prior to surgery. An equal number of specimens was collected from healthy volunteers, which were partners or accompanying persons of included patients. All samples from cases and controls originated from a Dutch population and were matched by sex and age and sample collection date (i.e. freezer storage duration) in both the discovery and validation cohort.[32] All patients in the discovery cohort were seen at the outpatient clinic of the Leiden University Medical Center between October 2002 and December 2008. For an independent validation cohort, blood specimens were collected between June 2016 and March 2018. All selected patients in the discovery and validation cohorts were candidates for curative surgery. However, not all patients underwent surgery due to preoperative metastases. PDAC diagnosis consisted of a combination of annual abdominal magnetic resonance imaging, magnetic resonance cholangiopancreatography and/or optionally endoscopic ultrasound. Furthermore, all surgical specimens were examined according to routine histological evaluation and the extent of the tumor spread was assessed by TNM classification.[33,34]

Blood samples in the validation cohort were obtained from twenty patients diagnosed with PDAC, two patients with duodenal and papillary carcinoma, two patients with neuroendocrine tumors and three patients with IPMN. A total of twenty-seven healthy controls were randomly selected from the LUMC Biobank. Cases and controls were matched by sex and age and sample collection date (i.e. freezer storage duration) in both the discovery and validation cohort.

This study was approved by the Medical Ethical Committee of the LUMC (protocol number P03-147). All patients and healthy volunteers provided written informed consent prior to blood collection.

Serum sample collection and plate design

Blood specimens from both the discovery and validation cohorts were collected and processed according to a standardized protocol.[35] Briefly, all blood samples were drawn by antecubital venipuncture. Approximately 8 mL of venous blood was collected in a 10 mL BD vacutainer SST II advance and centrifuged for 10 minutes at 1000×g. Processing of blood specimens took place within 4 hours after blood collection. After the centrifugation step serum samples were distributed into sterile, 500-µLbarcode-labeled aliquots and stored at -80°C until further analysis. Before measurements (i.e. serum *N*-glycome analysis) took place, each sample was aliquoted into 60 µL tubes.[35] One aliquot of each sample was then relocated into a 96-

well plate format according to a plate design, thus keeping cases and their age- and sex-matched controls on the same plate. Additionally, for technical quality control (QC) of the spectra, each plate contained a minimum of six in-house standards and two blanks.

Serum sample preparation and mass spectrometry analysis of glycans

N-glycans were enzymatically released from serum glycoproteins, chemically derivatized, purified, MS-analyzed, identified and quantified. Briefly, 6 μL of serum was used according to a previously reported protocol.[31] The global release of *N*-glycans was performed using the enzyme PNGase F (Roche Diagnostics, Mannheim, Germany). All following steps were carried out in a standardized manner on a Hamilton liquid handling platform. In a first step, all sialic acid residues at the non-reducing ends of the complex glycan structures were derivatized into stable end-products allowing the differentiation between α2,3- and α2,6-linked sialic acids by the introduced mass difference. Next, the glycans were purified using in-house developed cotton-based hydrophilic interaction liquid chromatography (HILIC) micro-tips. The purified glycans were eluted and premixed with sDHB matrix (5 mg/mL in 99% ACN with 1 mM NaOH). The mixture was spotted onto a MALDI target plate (800/384 MTP AnchorChip, Bruker Daltonics, Bremen, Germany) and spots were allowed to dry. Measurements were performed on a Bruker 15T solariX XR Fourier transform ion cyclotron resonance (FTICR)MS. The system was controlled by ftmsControl version 2.1.0 and spectra in an *m*/*z*-range from 1011.86 to 5000.00 were recorded with 1 M data points (i.e., transient length of 2.307 s). DataAnalysis Software 4.2 (Bruker Daltonics) was used for the visualization and data analysis of all MALDI-FTICR spectra. Sample preparation and subsequent glycan measurements were identical for all samples in both cohorts, however the validation cohort was processed five months after the discovery cohort.

Data processing and statistics

Serum *N*-glycan profiles were obtained from all 88 cases and 88 controls in the discovery cohort, of which 86 case-profiles and 84 control-profiles passed the quality criteria.[31] These profiles are further referred to as the discovery set. In the validation cohort, consisting of 27 cases and 27 controls, 26 case-profiles and 26 control-profiles passed. These profiles are further referred to as the validation set. For both the discovery and validation set, the same analyte list with 84 glycan compositions (Supporting information Table S-1) which passed the quality criteria[31] was used for data extraction with MassyTools version 0.1.8.1.31 To study general glycosylation features, such as fucosylation, branching, sialylation and bisection, derived traits were calculated to combine the effects of glycans with similar structures (Supporting

information Table S-2).

To evaluate the potential of total serum *N*-glycome analysis in differentiating PDAC patients from controls, logistic regression was performed for each glycoform individually as well as for each derived trait (Supporting information Table S-3 and S-4), using R version 3.3.2 (R Foundation for Statistical Computing, Vienna, Austria; Released 31 October 2016) and RStudio, version 1.0.136 (RStudio, Boston, MA; Released 21 December2016).[36] The odds ratios (ORs) were calculated with their 95% confidence intervals (CIs) assuming a Student's t-distribution and are referring to an increase of 1 SD in the tested traits. A fixed-effects model was used to combine the data of the discovery and validation set in a meta-analysis. Multiple testing correction (Bonferroni) was performed on the meta-analyzed data. In order to evaluate potential trait differences between the various cancer stages, stages Ia, Ib and IIa were merged into one sub-group, IIb was considered as a separate sub-group, and stages III and IV were also merged into one sub-group. For plotting purposes, the center line is median, box limits are upper and lower quartiles, and whiskers give the maximum and minimum values excluding any outliers. All points are individual measurements and outliers are the individual measurements larger than quartile 3 + 1.5xIQR or smaller than quartile1 - 1.5xIQR (IQR=interquartile range). For all glycan comparisons between case control subjects the significance level is stated in each corresponding plot after adjusting the p-value of Student's t-test using B-H method.

Receiver operating characteristic (ROC) analysis was performed by selecting derived traits representing the different glycosylation features that showed the strongest effect sizes (antennarity, fucosylation, sialylation) in the meta-analysis. Initially, five derived traits were used for the model, namely CA2 (diantennary species of complex glycans in spectrum), CA4(tetraantennary species of complex glycans in spectrum), A3FE (α2,6-sialylation of fucosylatedtriantennary glycans), A3F0L (α2,3-sialylation of nonfucosylatedtriantennary glycans) and CFa (antenna-fucosylation of complex glycans). Multiple combinations of these traits were then evaluated with regard to classification of diagnosis, resulting in a final model based on a combination of CA4, A3F0L and CFa. The model was trained using a randomly selected 75% of the discovery set and evaluated for its prediction value on the remaining 25% to prevent overfitting. More importantly, the prediction was replicated on the validationset. The power of the classification (area under the curve) was evaluated ten times with each time a new random selection of 75% of the discovery set, resulting in a mean power that was more robust than a single classification.

Results

The serum *N*-glycomes of PDAC patients and matched controls in a discovery and independent validation set (Table 1) were analyzed by mass spectrometry. Derived traits were calculated for structural features shared by multiple glycans, such as the level of antennarity (in the following abbreviated as CA), α2,3-linked sialylation (L), α2,6-linked sialylation (E), fucosylation (F) and bisection (B) (Figure 1). Data of consistent quality were obtained as assessed from 19 in-house standards that were included in the TSNG measurements. It is furthermore noted that the MS-based glycan profiles provide relative quantitative data that do not explain whether differences are caused by different serum protein concentrations or to which extent protein-specific glycosylation differences contribute. The data revealed age- and sex-associations of the glycomic signatures (Supporting information Figure S1) in accordance with literature[37] supporting the validity of the data. Logistic regression analysis was performed both at the single *N*-glycan level and the derived traits (Figure 1) revealing a total of 23 glycosylation features that where consistently found to differ between patients and controls as demonstrated by our meta-analysis (Table 2).

Table 1. Patient characteristics

	Discovery Set		**Validation Set**	
	Cases	**Controls**	**Cases**	**Controls**
	(n = 86)	**(n = 84)**	**(n = 26)**	**(n = 26)**
Female sex, n (%)	47 (54.7)	45 (53.4)	11 (42.3)	11 (42.3)
Age in years, mean (SD)	64.6 (11.1)	63.2 (10.0)	66.3 (10.5)	66.7 (5.4)
Diagnosis, n				
PDAC	86	n/a	20	n/a
Other	n/a	n/a	6	n/a
Stage, n				
Ia	6	n/a	0	n/a
Ib	8	n/a	1	n/a
IIa	10	n/a	3	n/a
IIb	41	n/a	7	n/a
III	3	n/a	1	n/a
IV	18	n/a	8	n/a

A strong increase in the antennarity of the glycans was found: tetraantennary *N*-glycans (CA4) were more abundant in PDAC profiles than in control samples (Table 2 and Figure 2), with a concomitant decrease in diantennary *N*-glycans (CA2;

Table 2 and Figure 2). Moreover, antenna-fucosylation (CFa, difucosylation) was increased (CFa;Table 2 and Figure 2). Both mono- and difucosylation (F and Fa, respectively) were increased for tri- and tetraantennary glycans (A3F, A4F, A3Fa, A4Fa; Table 2).

Table 2. Replicated meta-analyzed associations of serum *N*-glycans with pancreatic cancer.

Derived traits	Description of derived traits	Meta p-values	Odds ratio	Confidence interval
		Cases / Controls	Cases / Controls	Cases / Controls
Glycan type				
CA2	Diantennary species of complex glycans in spectrum	1.05E-08	0.35	(0.25-0.50)
CA4	Tetraantennary species of complex glycans in spectrum	9.21E-11	6.19	(3.57-10.75)
CFa	Antenna-fucosylation of complex glycans	2.31E-09	13.27	(5.68-30.98)
CB0	Non-bisected species of complex glycans in spectrum	5.12E-08	0.39	(0.27-0.54)
Fucosylation (F)				
A3F	Fucosylation of triantennary glycans	2.07E-07	2.34	(1.70-3.23)
A4F	Fucosylation of tetraantennary glycans	5.00E-06	2.04	(1.50-2.78)
A3Fa	Antenna-fucosylation of triantennary glycans	1.12E-08	5.35	(3.01-9.52)
A4Fa	Antenna-fucosylation of tetraantennary glycans	1.45E-06	2.45	(1.70-3.53)
A2LF	Fucosylation of α2,3-sialylated diantennary glycans	3.85E-08	2.67	(1.88-3.78)
A3LF	Fucosylation of α2,3-sialylated triantennary glycans	9.32E-09	2.68	(1.91-3.75)
A4LF	Fucosylation of α2,3-sialylated tetraantennary glycans	1.70E-06	2.13	(1.56-2.91)
A4EF	Fucosylation of α2,6-sialylated tetraantennary glycans	6.06E-06	2.03	(1.49-2.75)
Sialylation (S)				
A4F0S	Sialylation of nonfucosylated tetraantennary glycans	3.07E-06	0.48	(0.35-0.65)
A3FS	Sialylation of fucosylated triantennary glycans	7.41E-05	1.95	(1.40-2.71)
A4FS	Sialylation of fucosylated tetraantennary glycans	2.85E-06	2.1	(1.54-2.87)
α2,3-Linked sialylation (L)				
A2F0L	α2,3-sialylation of nonfucosylated diantennary glycans	5.74E-07	0.38	(0.26-0.55)
A3F0L	α2,3-sialylation of nonfucosylated triantennary glycans	8.34E-09	0.34	(0.23-0.49)
A4F0L	α2,3-sialylation of nonfucosylated tetraantennary glycans	4.50E-07	0.44	(0.32-0.61)
α2,6-Linked sialylation (E)				
A3E	α2,6-sialylation of triantennary glycans	3.22E-07	2.41	(1.72-3.38)
A2F0E	α2,6-sialylation of nonfucosylated diantennary glycans	3.18E-07	2.44	(1.73-3.43)
A3F0E	α2,6-sialylation of nonfucosylated triantennary glycans	2.62E-09	3.5	(2.32-5.30)
A3FE	α2,6-sialylation of fucosylated triantennary glycans	1.20E-09	3.99	(2.55-6.24)
A4FE	α2,6-sialylation of fucosylated tetraantennary glycans	2.02E-08	2.63	(1.88-3.69)

NOTE. Dark grey and light gray shading indicate positive and negative associations, respectively, with the healthy controls being the reference. See Supporting information Table S-3 for the complete list of tests performed.

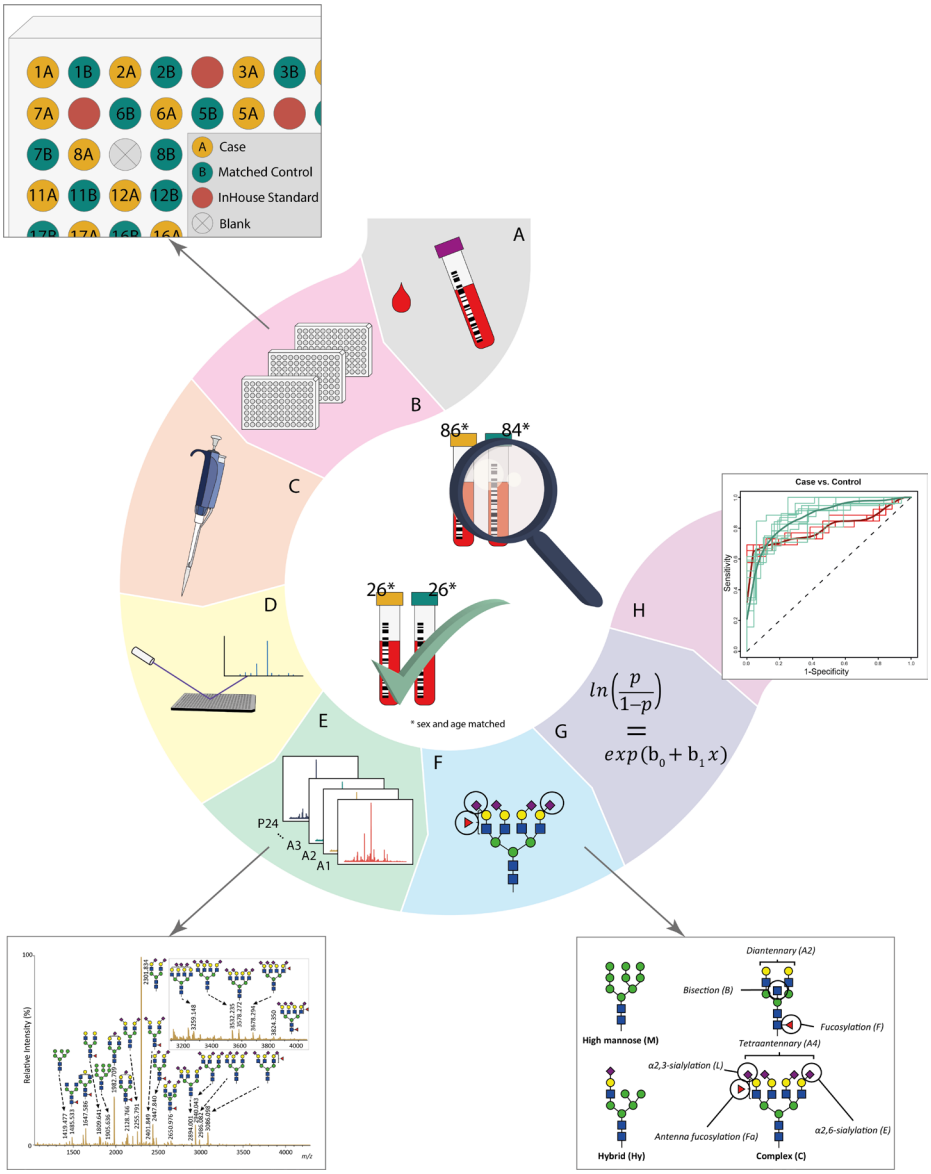


Figure 1. Workflow of *N*-glycosylation analysis of discovery and validation pancreatic cancer case-control cohorts for classification analysis. (A) Collection of serum samples from pancreatic cancer patients and healthy controls. (B) Random distribution of age- and sex-matched case-control pairs, in-house standards and blanks. (C) Automated sample preparation including enzymatic glycan release, derivatization and purification. (D) MALDI-FTICR-MS analysis of *N*-glycome. (E) MS-spectrum preprocessing, annotation and quality control. (F) Derived trait calculation for the analysis of glycosylation features. (G) Logistic regression analysis of both cohorts, followed by meta-analysis of the data. (H) ROC-analysis to test glycosylation traits for their classification power.

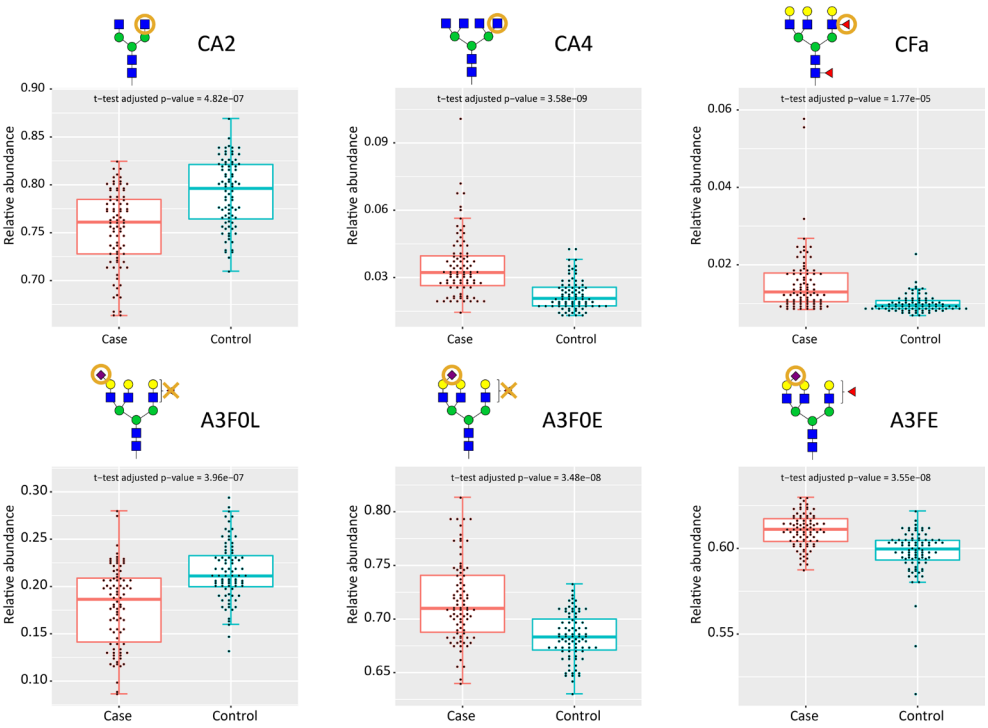


Figure 2. Main replicated associations between *N*-glycan traits and pancreatic cancer, based on the data from the discovery cohort with corresponding Student's t-test adjusted p-values.

Also, an increase in α2,6-linked sialylation and a decrease in α2,3-linked sialylation was observed(A2F0L, A3F0L, A4F0L; Table 2). Increased α2,6-linked sialylation was observed in both fucosylated and non-fucosylated di-, tri- and tetraantennary glycans (e.g. A3E, A3F0E, A3FE; Table 2).We further evaluated whether glycan derived traits were associated with cancer stages such as depicted in Table 1, but no differences were found between the various stages (details explained in the Methods section).

Finally, receiver operating characteristic (ROC) curves were calculated for selected glycan traits. The resulting ROC curve illustrates the power of differentiating PDAC from matched control samples (Figure 3). With an AUC of 0.88 the discriminative performance of the discovery set was good. At the optimal case probability score cut-off, the sensitivity and specificity were 0.85 and 0.71, respectively. The signature was replicated in an independent validation cohort with a good AUC of 0.81, and with a sensitivity and specificity of 0.72 and 0.75, respectively.

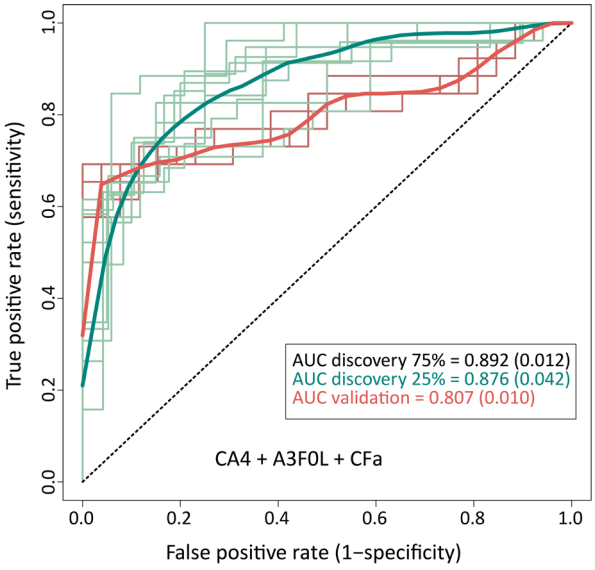


Figure 3. ROC-analysis with a model based on CA4, A3F0L and CFa. The model was trained with a random selection of 75% of the spectra in the discovery cohort and applied to the remaining 25% of the cohort to test for its prediction value. Moreover, it was applied to an independent validation cohort to test for its classification power. This analysis was repeated ten times, to increase the robustness of AUCs. The means (and SDs) of 10 predictions are reported for the respective AUC.

Discussion

The objective of this study was to explore the potential of serum protein *N*-glycan profiles from PDAC patients to serve as a biomarker panel, aiming for the development of a blood-based test for diagnosis of PDAC. Using our recently established analytical glycomics platform, 112 patient sera were analyzed and compared to 110 healthy control samples. The observed *N*-glycosylation changes in the discovery cohort of PDAC patients were replicated in an independent validation cohort.

Major glycosylation differences were found between PDAC patients and controls for *N*-glycan antennarity, fucosylation and sialylation. Notably, with regard to sialylation, our approach included an evaluation of α2,3- and α2,6-linked sialic acids separately. We found twenty-three glycosylation traits to be associated with PDAC in a meta-analysis of the two sample sets.

PDAC patients showed higher α2,6-linked sialylation than controls. From tumor cell surface analysis it is known that an increase in overall α2,6-linked sialic acids associates with cancer progression.[38,39] A possible explanation is that α2,6-linked sialic acids promote cancer cell survival since binding of pro-apoptotic galectins to cell surface glycans is blocked by these structures. In contrast, α2,3-linked sialic acids do not inhibit galectin binding.[38,39] With regard to protein N-glycosylation,similar functions of α2,6-linked sialylation have been suggested, however the mechanisms in this case are not yet understood.[39]

The current study also demonstrated elevated levels of tri- and tetraantennary *N*-glycans and a concomitant decrease of diantennary *N*-glycans in PDAC patients. Previously, similar observations in PDAC patients have been reported with regard to branching in total serum glycosylation profiles as well as in studies on specific glycoproteins, such as α-1-acid glycoprotein (AGP) and haptoglobin.[24–26,40] The first study reported elevated levels of tri- and tetraantennary glycans, however with a limited sample set of two pairs of cancer and normal samples only.[24] The latter studies reported increased branching of AGP-derived glycans with limited sample sets of 19 PDAC patients, 6 chronic pancreatitis patients and 6 controls, and increased branching in HPT and transferrin.[25,26] Increased tri- and tetraantennary glycans have furthermore been reported in association with progression of disease in sera and cell lines from PDAC patients.[24,27,28] With regard to other cancer types,an increase in branching has been observed in brain and colorectal cancer.[21,41–43]

Besides increased branching, an increased fucosylation of tri- and tetraantennary *N*-glycans was found in PDAC patients. Increased fucosylation has been reported in various types of cancer such as hepatocellular carcinoma, oral and colorectal cancer.[21,44,45] Also the previously mentioned glycoprotein studies on

AGP and HPT reported an increased fucosylation.[25,26] Interestingly, Akimoto and co-workers studied serum *N*-glycan profiles of 79 patients with IPMN and found a potential marker for invasive IPMNs based on an increased expression of fucosylated complex-type glycans. Unfortunately, the *N*-glycan profiles in this study were not compared to those obtained from healthy control individuals.[40]

We found an increase of fucosylation of triantennary and tetraantennary glycans (A3F and A4F) in PDAC patients, specifically, in glycans containing α2,3-linked sialic acids (A2LF, A3LF and A4LF). The combination of α2,3-linked sialylation with fucosylation suggests the formation of sialyl-Lewis X (sLeX) moieties. The increase of sLeX expression has been reported with regard to pancreatic cancer.[26,28] In addition, increased sLeX expression on AGP and HPT has been linked to various cancers (e.g. pancreatic cancer, lung cancer, advanced ovarian cancer and prostate cancer)[18,25,46–49] and chronic inflammation (e.g. rheumatoid arthritis and inflammatory bowel disease).[50] As discussed in the Introduction, the marker CA19-9 is based on a sialyl-Lewis A (sLeA) epitope. Although this structure differs from sLeX with regard to glycosidic linkages, both point toward the importance of sialylation.

Increased branching and sLeX expression have been found in acute phase proteins which are released by the liver in the event of cancer, but also in case of infection, surgery and inflammatory conditions.[50] The relation between inflammation and cancer has been discussed in a review, with the hypothesis that both glycosylation changes are a systemic side effect of inflammatory cytokines stimulating the liver under influence of the tumor.[51] It has been demonstrated that the tumor microenvironment contains large amounts of these cytokines and that inflammatory pathways are involved in the development of tumors.[52] The expression of these cytokines was also confirmed in studies on cell lines and tissues.[53]

The here applied glycomics workflow is specifically suited for a high-throughput and relatively fast "cancer glycosylation profiling" of body fluids and cell or tissue material from a clinical cohort. This strategy does not provide detailed information on the protein origin of the potential glycan markers. This limitation is well-known in the glycobiology community and can be tackled with in-depth glycoproteomic analyses that come with their own challenges. Thus, although the analysis of total serum *N*-glycosylation shows strong associations with PDAC, it is expected that analysis of specific glycoproteins might further improve accuracy.

The need for a screening test for pancreatic cancer is high, especially for patients with increased inherited risk. A screening test should meet specific requirements and should exhibit suitable sensitivity and specificity specifications.[29] The current results are promising in terms of the discriminative performance for sensitivity and specificity, but translation into the clinic depends on the application. The screening

of patients with a genetically increased risk for PDAC would be a first step since no test is currently available to support clinicians. For the general population the discriminative performance found in this study might be insufficient for application but could possibly be complementary to the CA 19-9 test.[14] Detection of PDAC at an earlier stage needs further investigation, since early detection is an important argument for population screening.[6]

The discriminating performance of case-control ROC-analysis was good, indicating a strong difference in *N*-glycosylation profiles of PDAC patients and healthy controls. However, as indicated above, the *N*-glycosylation shift we found in PDAC patients is not necessarily specific for pancreatic cancer. In this study, only PDAC patients and healthy volunteers were included, while in a clinical application other diseases might interfere with the determination of the PDAC cases. To address the specificity of the discriminating signals in this study, future research should compare PDAC signatures with those of benign diseases (e.g. pancreatitis) as well as other types of cancers and inflammatory diseases.[30]

Conclusions

In this study, serum *N*-glycome analysis with sialic acid isomer differentiation and ultrahigh resolution MS was performed to classify PDAC patients from healthy controls. Three major *N*-glycosylation differences were observed and validated between cases and healthy controls, namely (antenna-) fucosylation of complex glycans, branching of complex glycans and increased α2,6-linked sialylation compared to the α2,3-linked analogues. Combination of various *N*-glycosylation traits resulted in classification performance that can function as a target for follow-up glycomics research aiming for development of a blood-based clinical test. In future research the specificity of the observed changes needs to be addressed by including samples from benign pancreatic diseases including inflammation and preferably other cancer types. In addition, longitudinal analysis is warranted to determine the potential for early detection based on the here reported serum *N*-glycan disease signatures.

Acknowledgements

The authors thank Ronald van Vlierberghe and Priscilla van der Wilk at the laboratory of surgery for their technical assistance and dr. Yassene Mohammed at the Center for Proteomics and Metabolomics for his assistance with the ROC-analysis.

References

(1) (1) Weinrich, M.; Bochow, J.; Kutsch, A. L.; Alsfasser, G.; Weiss, C.; Klar, E.; Rau, B. M. High Compliance with Guideline Recommendations but Low Completion Rates of Adjuvant Chemotherapy in Resected Pancreatic Cancer: A Cohort Study. Ann. Med. Surg. 2018, 32 (June), 32–37. https://doi.org/10.1016/j.amsu.2018.06.004.

(2) Kardosh, A.; Lichtensztajn, D. Y.; Gubens, M. A.; Kunz, P. L.; Fisher, G. A.; Clarke, C. A. Long-Term Survivors of Pancreatic Cancer: A California Population-Based Study. Pancreas 2018, 47 (8), 958–966. https://doi.org/10.1097/MPA.0000000000001133.

(3) Raman, P.; Maddipati, R.; Lim, K. H.; Tozeren, A. Pancreatic Cancer Survival Analysis Defines a Signature That Predicts Outcome. PLoS One 2018, 13 (8), 1–18. https://doi.org/10.1371/journal.pone.0201751.

(4) Navarro, E. B.; López, E. V.; Quijano, Y.; Caruso, R.; Ferri, V.; Durand, H.; Cabrera, I. F.; Reques, E. D.; Ielpo, B.; Glagolieva, A. Y.; Plaza, C. Impact of BRCA1/2 Gene Mutations on Survival of Patients with Pancreatic Cancer: A Case-Series Analysis. Ann. Hepato-Biliary-Pancreatic Surg. 2019, 23 (2), 200. https://doi.org/10.14701/ahbps.2019.23.2.200.

(5) Lennon, A. M.; Wolfgang, C. L.; Canto, M. I.; Klein, A. P.; Herman, J. M.; Goggins, M.; Fishman, E. K.; Kamel, I.; Weiss, M. J.; Diaz, L. A.; Papadopoulos, N.; Kinzler, K. W.; Vogelstein, B.; Hruban, R. H. The Early Detection of Pancreatic Cancer: What Will It Take to Diagnose and Treat Curable Pancreatic Neoplasia? Cancer Res. 2014, 74 (13), 3381–3389. https://doi.org/10.1158/0008-5472.CAN-14-0734.

(6) Singhi, A. D.; Koay, E. J.; Chari, S. T.; Maitra, A. Early Detection of Pancreatic Cancer: Opportunities and Challenges. Gastroenterology 2019, 156 (7), 2024–2040. https://doi.org/10.1053/j.gastro.2019.01.259.

(7) Vasen, H.; Ibrahim, I.; Robbers, K.; Van Mil, A. M.; Potjer, T.; Bonsing, B. A.; Bergman, W.; Wasser, M.; Morreau, H.; De Vos Tot Nederveen Cappel, W. H.; Ponce, C. G.; Carrato, A.; Earl, J.; Mocci, E.; Vazquez-Sequeiros, E.; Sanjuanbenito, A.; Muñoz-Beltran, M.; Montans, J.; Slater, E. P.; Matthäi, E.; Fendrich, V.; Bartsch, D. K.; Schicker, C.; Steinkamp, M.; Figiel, J.; Klöppel, G.; Langer, P.; Esposito, I. Benefit of Surveillance for Pancreatic Cancer in High-Risk Individuals: Outcome of Long-Term Prospective Follow-up Studies from Three European Expert Centers. J. Clin. Oncol. 2016, 34 (17), 2010–2019. https://doi.org/10.1200/JCO.2015.64.0730.

(8) McAllister, F.; Montiel, M. F.; Uberoi, G. S.; Uberoi, A. S.; Maitra, A.; Bhutani, M. S. Current Status and Future Directions for Screening Patients at High Risk for Pancreatic Cancer. Gastroenterol. Hepatol. (N. Y). 2017, 13 (5), 268–275.

(9) Ferlay, J.; Partensky, C.; Bray, F. More Deaths from Pancreatic Cancer than Breast Cancer in the EU by 2017. Acta Oncol. (Madr). 2016, 55 (9–10), 1158–1160. https://doi.org/10.1080/0284186X.2016.1197419.

(10) Chu, L. C.; Goggins, M. G.; Fishman, E. K. Diagnosis and Detection of Pancreatic Cancer. Cancer J. (United States) 2017, 23 (6), 333–342. https://doi.org/10.1097/PPO.0000000000000290.

(11) Yamashita, Y.-I.; Okabe, H.; Hayashi, H.; Imai, K.; Nakagawa, S.; Nakao, Y.; Yusa, T.; Itoyama, R.; Yama, T.; Umesaki, N.; Arima, K.; Miyata, T.; Chikamoto, A.; Baba, H. Usefulness of 18-FDG PET/CT in Detecting Malignancy in Intraductal Papillary Mucinous Neoplasms of the Pancreas. Anticancer Res. 2019, 39 (5), 2493–2499. https://doi.org/10.21873/anticanres.13369.

(12) Dickerson, L. D.; Farooq, A.; Bano, F.; Kleeff, J.; Baron, R.; Raraty, M.; Ghaneh, P.; Sutton, R.; Whelan, P.; Campbell, F.; Healey, P.; Neoptolemos, J. P.; Yip, V. S. Differentiation of Autoimmune Pancreatitis from Pancreatic Cancer Remains Challenging. World J. Surg. 2019, 43 (6), 1604–1611. https://doi.org/10.1007/s00268-019-04928-w.

(13) Anderson, L. Within Sight of a Rational Pipeline for Development of Protein Diagnostics. Clin. Chem. 2012, 58 (1), 28–30. https://doi.org/10.1373/clinchem.2011.173377.

(14) Ansari, D.; Chen, B. C.; Dong, L.; Zhou, M. T.; Andersson, R. Pancreatic Cancer: Translational Research Aspects and Clinical Implications. World J. Gastroenterol. 2012, 18 (13), 1417–1424. https://doi.org/10.3748/wjg.v18.i13.1417.

(15) Wright, I.; Van Eyk, J. E. A Roadmap to Successful Clinical Proteomics. Clin. Chem. 2017, 63 (1), 245–247. https://doi.org/10.1373/clinchem.2016.254664.

(16) Smith, L. M.; Kelleher, N. L. Proteoforms as the next Proteomics Currency. Science (80-.). 2018, 359 (6380), 1106–1107. https://doi.org/10.1126/science.aat1884.

(17) Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; Ge, Y.; Gunawardena, J.; Hendrickson, R. C.; Hergenrother, P. J.; Huber, C. G.; Ivanov, A. R.; Jensen, O. N.; Jewett, M. C.; Kelleher, N. L.; Kiessling, L. L.; Krogan, N. J.; Larsen, M. R.; Loo, J. A.; Ogorzalek Loo, R. R.; Lundberg, E.; Maccoss, M. J.; Mallick, P.; Mootha, V. K.; Mrksich, M.; Muir, T. W.; Patrie, S. M.; Pesavento, J. J.; Pitteri, S. J.; Rodriguez, H.; Saghatelian, A.; Sandoval, W.; Schlüter, H.; Sechi, S.; Slavoff, S. A.;

Smith, L. M.; Snyder, M. P.; Thomas, P. M.; Uhlén, M.; Van Eyk, J. E.; Vidal, M.; Walt, D. R.; White, F. M.; Williams, E. R.; Wohlschlager, T.; Wysocki, V. H.; Yates, N. A.; Young, N. L.; Zhang, B. How Many Human Proteoforms Are There? Nat. Chem. Biol. 2018, 14 (3), 206–214. https://doi.org/10.1038/nchembio.2576.

(18) Saldova, R.; Wormald, M. R.; Dwek, R. A.; Rudd, P. M. Glycosylation Changes on Serum Glycoproteins in Ovarian Cancer May Contribute to Disease Pathogenesis. Dis. Markers 2008, 25 (4–5), 219–232. https://doi.org/10.1155/2008/601583.

(19) Alley, W. R.; Madera, M.; Mechref, Y.; Novotny, M. V. Chip-Based Reversed-Phase Liquid Chromatography-Mass Spectrometry of Permethylated N-Linked Glycans: A Potential Methodology for Cancer-Biomarker Discovery. Anal. Chem. 2010, 82 (12), 5095–5106. https://doi.org/10.1021/ac100131e.

(20) Reiding, K. R.; Bondt, A.; Hennig, R.; Gardner, R. A.; O'Flaherty, R.; Trbojević-Akmačić, I.; Shubhakar, A.; Hazes, J. M. W.; Reichl, U.; Fernandes, D. L.; Pučić-Baković, M.; Rapp, E.; Spencer, D. I. R.; Dolhain, R. J. E. M.; Rudd, P. M.; Lauc, G.; Wuhrer, M. High-Throughput Serum N -Glycomics: Method Comparison and Application to Study Rheumatoid Arthritis and Pregnancy-Associated Changes. Mol. Cell. Proteomics 2019, 18 (1), 3–15. https://doi.org/10.1074/mcp.RA117.000454.

(21) de Vroome, S. W.; Holst, S.; Girondo, M. R.; van der Burgt, Y. E. M.; Mesker, W. E.; Tollenaar, R. A. E. M.; Wuhrer, M. Serum N-Glycome Alterations in Colorectal Cancer Associate with Survival. Oncotarget 2018, 9 (55), 30610–30623. https://doi.org/10.18632/oncotarget.25753.

(22) Geyer, H.; Geyer, R. Strategies for Analysis of Glycoprotein Glycosylation. Biochim. Biophys. Acta - Proteins Proteomics 2006, 1764 (12), 1853–1869. https://doi.org/10.1016/j.bbapap.2006.10.007.

(23) Cummings, R. D. The Repertoire of Glycan Determinants in the Human Glycome. Mol. Biosyst. 2009, 5 (10), 1087. https://doi.org/10.1039/b907931a.

(24) Zhao, J.; Qiu, W.; Simeone, D. M.; Lubman, D. M. N-Linked Glycosylation Profiling of Pancreatic Cancer Serum Using Capillary Liquid Phase Separation Coupled with Mass Spectrometric Analysis. J. Proteome Res. 2007, 6 (3), 1126–1138. https://doi.org/10.1021/pr0604458.

(25) Balmaña, M.; Giménez, E.; Puerta, A.; Llop, E.; Figueras, J.; Fort, E.; Sanz-Nebot, V.; de Bolós, C.; Rizzi, A.; Barrabés, S.; de Frutos, M.; Peracaula, R. Increased A1-3 Fucosylation of α-1-Acid Glycoprotein (AGP) in Pancreatic Cancer. J. Proteomics 2016, 132, 144–154. https://doi.org/10.1016/j.jprot.2015.11.006.

(26) Sarrats, A.; Saldova, R.; Pla, E.; Fort, E.; Harvey, D. J.; Struwe, W. B.; de Llorens, R.; Rudd, P. M.; Peracaula, R. Glycosylation of Liver Acute-Phase Proteins in Pancreatic Cancer and Chronic Pancreatitis. Proteomics - Clin. Appl. 2010, 4 (4), 432–448. https://doi.org/10.1002/prca.200900150.

(27) Park, H.; Hwang, M. P.; Kim, Y.; Kim, K. J.; Jin, J. M.; Kim, Y. H.; Yang, Y. H.; Lee, K. H.; Kim, Y. G. Mass Spectrometry-Based N-Linked Glycomic Profiling as a Means for Tracking Pancreatic Cancer Metastasis. Carbohydr. Res. 2015, 413 (1), 5–11. https://doi.org/10.1016/j.carres.2015.04.019.

(28) Munkley, J. The Glycosylation Landscape of Pancreatic Cancer (Review). Oncol. Lett. 2019, 17 (3), 2569–2575. https://doi.org/10.3892/ol.2019.9885.

(29) Kailemia, M. J.; Park, D.; Lebrilla, C. B. Glycans and Glycoproteins as Specific Biomarkers for Cancer. Anal. Bioanal. Chem. 2017, 409 (2), 395–410. https://doi.org/10.1007/s00216-016-9880-6.

(30) Krishnan, S.; Whitwell, H. J.; Cuenco, J.; Gentry-Maharaj, A.; Menon, U.; Pereira, S. P.; Gaspari, M.; Timms, J. F. Evidence of Altered Glycosylation of Serum Proteins Prior to Pancreatic Cancer Diagnosis. Int. J. Mol. Sci. 2017, 18 (12), 8–10. https://doi.org/10.3390/ijms18122670.

(31) Vreeker, G. C. M.; Nicolardi, S.; Bladergroen, M. R.; van der Plas, C. J.; Mesker, W. E.; Tollenaar, R. A. E. M.; van der Burgt, Y. E. M.; Wuhrer, M. Automated Plasma Glycomics with Linkage-Specific Sialic Acid Esterification and Ultrahigh Resolution MS. Anal. Chem. 2018, 90 (20), 11955–11961. https://doi.org/10.1021/acs.analchem.8b02391.

(32) Ren, S.; Zhang, Z.; Xu, C.; Guo, L.; Lu, R.; Sun, Y.; Guo, J.; Qin, R.; Qin, W.; Gu, J. Distribution of IgG Galactosylation as a Promising Biomarker for Cancer Screening in Multiple Cancer Types. Cell Res. 2016, 26 (8), 963–966. https://doi.org/10.1038/cr.2016.83.

(33) Lu, J. P.; Knežević, A.; Wang, Y. X.; Rudan, I.; Campbell, H.; Zou, Z. K.; Lan, J.; Lai, Q. X.; Wu, J. J.; He, Y.; Song, M. S.; Zhang, L.; Lauc, G.; Wang, W. Screening Novel Biomarkers for Metabolic Syndrome by Profiling Human Plasma N-Glycans in Chinese Han and Croatian Populations. J. Proteome Res. 2011, 10 (11), 4959–4969. https://doi.org/10.1021/pr2004067.

(34) Vasen, H. F. A.; Wasser, M.; Van Mil, A.; Tollenaar, R. A.; Konstantinovski, M.; Gruis, N. A.; Bergman, W.; Hes, F. J.; Hommes, D. W.; Offerhaus, G. J. A.; Morreau, H.; Bonsing, B. A.; De Vos Tot Nederveen Cappel, W. H. Magnetic Resonance Imaging Surveillance Detects Early-Stage Pancreatic Cancer in Carriers of a P16-Leiden Mutation. Gastroenterology 2011, 140 (3), 850–856. https://doi.org/10.1053/j.gastro.2010.11.048.

(35) Velstra, B.; Bonsing, B. A.; Mertens, B. J.; van der Burgt, Y. E. M.; Huijbers, A.; Vasen, H.; Mesker, W. E.; Deelder, A.

M.; Tollenaar, R. A. E. M. Detection of Pancreatic Cancer Using Serum Protein Profiling. HPB 2013, 15 (8), 602–610. https://doi.org/10.1111/hpb.12017.

(36) Chong, J.; Xia, J. MetaboAnalystR: An R Package for Flexible and Reproducible Analysis of Metabolomics Data. Bioinformatics 2018, 34 (24), 4313–4314. https://doi.org/10.1093/bioinformatics/bty528.

(37) Reiding, K. R.; Ruhaak, L. R.; Uh, H. W.; El Bouhaddani, S.; Van Den Akker, E. B.; Plomp, R.; McDonnell, L. A.; Houwing-Duistermaat, J. J.; Slagboom, P. E.; Beekman, M.; Wuhrer, M. Human Plasma N-Glycosylation as Analyzed by Matrix-Assisted Laser Desorption/Ionization-Fourier Transform Ion Cyclotron Resonance-MS Associates with Markers of Inflammation and Metabolic Health. Mol. Cell. Proteomics 2017, 16 (2), 228–242. https://doi.org/10.1074/mcp.M116.065250.

(38) Schultz, M. J.; Swindall, A. F.; Bellis, S. L. Regulation of the Metastatic Cell Phenotype by Sialylated Glycans. Cancer Metastasis Rev. 2012, 31 (3–4), 501–518. https://doi.org/10.1007/s10555-012-9359-7.

(39) Lu, J.; Gu, J. Significance of β-Galactoside A2,6 Sialyltranferase 1 in Cancers. Molecules 2015, 20 (5), 7509–7527. https://doi.org/10.3390/molecules20057509.

(40) Akimoto, Y.; Nouso, K.; Kato, H.; Miyahara, K.; Dohi, C.; Morimoto, Y.; Kinugasa, H.; Tomoda, T.; Yamamoto, N.; Tsutsumi, K.; Kuwaki, K.; Onishi, H.; Ikeda, F.; Nakamura, S.; Shiraha, H.; Takaki, A.; Okada, H.; Amano, M.; Nishimura, S. I.; Yamamoto, K. Serum N-Glycan Profiles in Patients with Intraductal Papillary Mucinous Neoplasms of the Pancreas. Pancreatology 2015, 15 (4), 432–438. https://doi.org/10.1016/j.pan.2015.05.470.

(41) Dennis, J. W.; Laferté, S.; Waghorne, C.; Breitman, M. L.; Kerbel, R. S. B1-6 Branching of Asn-Linked Oligosaccharides Is Directly Associated with Metastasis. Science (80-.). 1987, 236 (4801), 582–585. https://doi.org/10.1126/science.2953071.

(42) Kizuka, Y.; Taniguchi, N. Enzymes for N-Glycan Branching and Their Genetic and Nongenetic Regulation in Cancer. Biomolecules 2016, 6 (2), 1–21. https://doi.org/10.3390/biom6020025.

(43) Zhao, Y.; Sato, Y.; Isaji, T.; Fukuda, T.; Matsumoto, A.; Miyoshi, E.; Gu, J.; Taniguchi, N. Branched N-Glycans Regulate the Biological Functions of Integrins and Cadherins. FEBS J. 2008, 275 (9), 1939–1948. https://doi.org/10.1111/j.1742-4658.2008.06346.x.

(44) Comunale, M. A.; Rodemich-Betesh, L.; Hafner, J.; Wang, M.; Norton, P.; di Bisceglie, A. M.; Block, T.; Mehta, A. Linkage Specific Fucosylation of Alpha-1-Antitrypsin in Liver Cirrhosis and Cancer Patients: Implications for a Biomarker of Hepatocellular Carcinoma. PLoS One 2010, 5 (8), 1–9. https://doi.org/10.1371/journal.pone.0012419.

(45) Shah, M.; Telang, S.; Raval, G.; Shah, P.; Patel, P. S. Serum Fucosylation Changes in Oral Cancer and Oral Precancerous Conditions: α-L-Fucosidase as a Marker. Cancer 2008, 113 (2), 336–346. https://doi.org/10.1002/cncr.23556.

(46) Kossowska, B.; Ferens-Sieczkowska, M.; Gancarz, R.; Passowicz-Muszyńska, E.; Jankowska, R. Fucosylation of Serum Glycoproteins in Lung Cancer Patients. Clin. Chem. Lab. Med. 2005, 43 (4), 361–369. https://doi.org/10.1515/CCLM.2005.066.

(47) Fujimura, T.; Shinohara, Y.; Tissot, B.; Pang, P. C.; Kurogochi, M.; Saito, S.; Arai, Y.; Sadilek, M.; Murayama, K.; Dell, A.; Nishimura, S. I.; Hakomori, S. I. Glycosylation Status of Haptoglobin in Sera of Patients with Prostate Cancer vs. Benign Prostate Disease or Normal Subjects. Int. J. Cancer 2008, 122 (1), 39–49. https://doi.org/10.1002/ijc.22958.

(48) Ishikawa, T.; Yoneyama, T.; Tobisawa, Y.; Hatakeyama, S.; Kurosawa, T.; Nakamura, K.; Narita, S.; Mitsuzuka, K.; Duivenvoorden, W.; Pinthus, J. H.; Hashimoto, Y.; Koie, T.; Habuchi, T.; Arai, Y.; Ohyama, C. An Automated Micro-Total Immunoassay System for Measuring Cancer-Associated A2,3-Linked Sialyl N-Glycan-Carrying Prostate-Specific Antigen May Improve the Accuracy of Prostate Cancer Diagnosis. Int. J. Mol. Sci. 2017, 18 (2). https://doi.org/10.3390/ijms18020470.

(49) Matsumoto, T.; Hatakeyama, S.; Yoneyama, T.; Tobisawa, Y.; Ishibashi, Y.; Yamamoto, H.; Yoneyama, T.; Hashimoto, Y.; Ito, H.; Nishimura, S. I.; Ohyama, C. Serum N-Glycan Profiling Is a Potential Biomarker for Castration-Resistant Prostate Cancer. Sci. Rep. 2019, 9 (1), 1–8. https://doi.org/10.1038/s41598-019-53384-y.

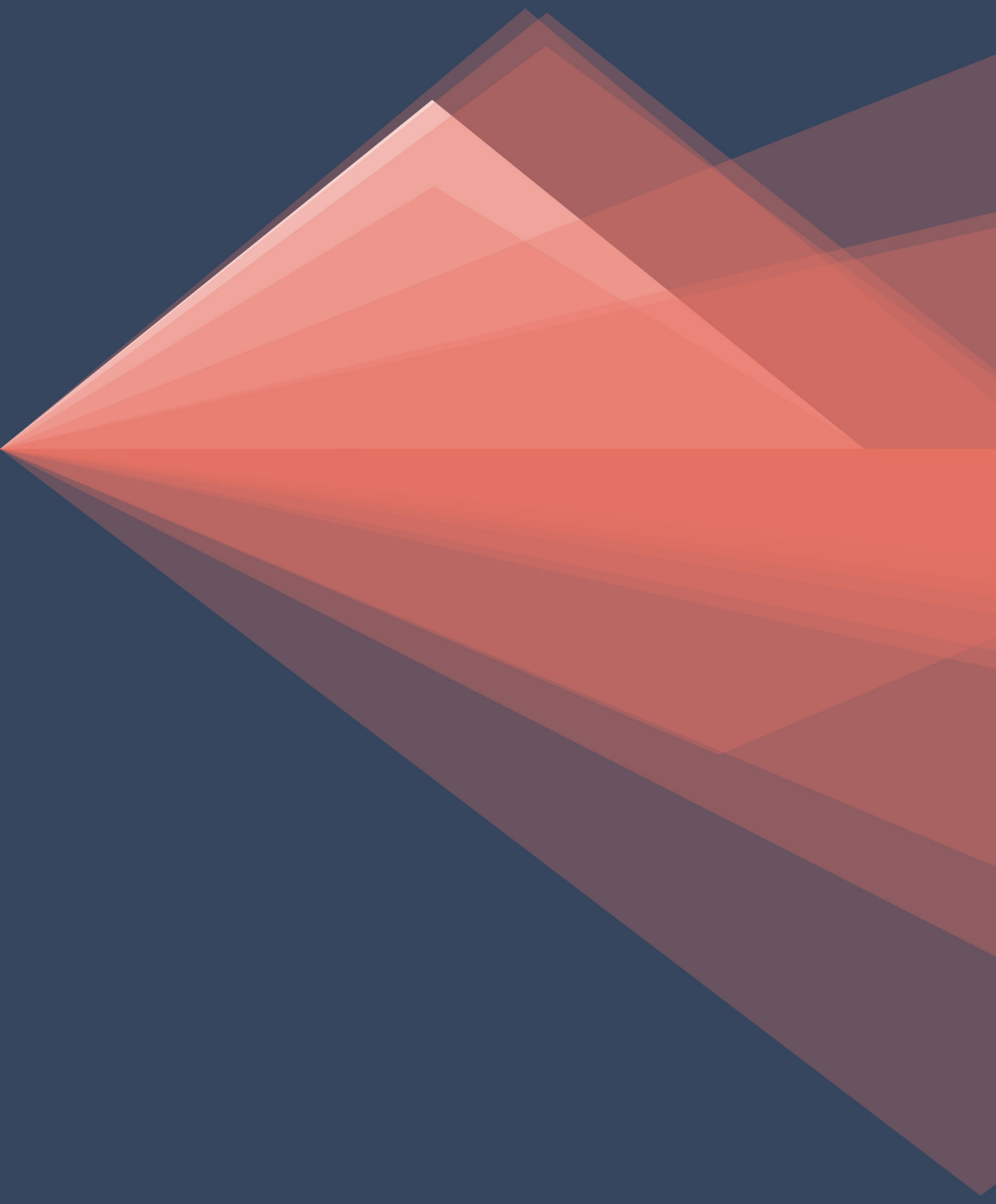
(50) Higai, K.; Aoki, Y.; Azuma, Y.; Matsumoto, K. Glycosylation of Site-Specific Glycans of A1-Acid Glycoprotein and Alterations in Acute and Chronic Inflammation. Biochim. Biophys. Acta - Gen. Subj. 2005, 1725 (1), 128–135. https://doi.org/10.1016/j.bbagen.2005.03.012.

(51) Arnold, J. N.; Saldova, R.; Abd Hamid, U. M.; Rudd, P. M. Evaluation of the Serum N-Linked Glycome for the Diagnosis of Cancer and Chronic Inflammation. Proteomics 2008, 8 (16), 3284–3293. https://doi.org/10.1002/pmic.200800163.

(52) Mantovani, A.; Allavena, P.; Sica, A.; Balkwill, F. Cancer-Related Inflammation. Nature 2008, 454 (7203), 436–444. https://doi.org/10.1038/nature07205.

(53) Bellone, G.; Smirne, C.; Mauri, F. A.; Tonel, E.; Carbone, A.; Buffolino, A.; Dughera, L.; Robecchi, A.; Pirisi, M.;

Emanuelli, G. Cytokine Expression Profile in Human Pancreatic Carcinoma Cells and in Surgical Specimens: Implications for Survival. Cancer Immunol. Immunother. 2006, 55 (6), 684–698. https://doi.org/10.1007/s00262-005-0047-0.



CHAPTER 6

Serum *N*-glycan profiles differ for various breast cancer subtypes

Based on:

Vreeker, G.C.M.[1,2], *Vangangelt, K.M.H.*[1], *Bladergroen, M.R.*[2], *Nicolardi, S.*[1,2], *Mesker, W.E.*[1], *Wuhrer, M.*[2], *Van der Burgt, Y.E.M.*[2], *Tollenaar, R.A.E.M.*[1]*: Serum N-glycan profiles differ for various breast cancer subtypes. Glycoconj. J. (2021). doi:10.1007/s10719-021-10001-3. Supplemental info available.*

1 Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands

2 Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands

Abstract

Breast cancer is the most prevalent cancer in women. Early detection of this disease improves survival and therefore population screenings, based on mammography, are performed. However, the sensitivity of this screening modality is not optimal and new screening methods, such as blood tests, are being explored. Most of the analyses that aim for early detection focus on proteins in the bloodstream.

In this study, the biomarker potential of total serum *N*-glycosylation analysis was explored with regard to detection of breast cancer. In an age-matched case-control setup serum protein *N*-glycan profiles from 145 breast cancer patients were compared to those from 171 healthy individuals. *N*-glycans were enzymatically released, chemically derivatized to preserve linkage-specificity of sialic acids and characterized by high resolution mass spectrometry. Logistic regression analysis was used to evaluate associations of specific *N*-glycan structures as well as *N*-glycosylation traits with breast cancer.

In a case-control comparison three associations were found, namely a lower level of a two triantennary glycans and a higher level of one tetraantennary glycan in cancer patients. Of note, various other *N*-glycomic signatures that had previously been reported were not replicated in the current cohort. It was further evaluated whether the lack of replication of breast cancer *N*-glycomic signatures could be partly explained by the heterogenous character of the disease since the studies performed so far were based on cohorts that included diverging subtypes in different numbers. It was found that serum *N*-glycan profiles differed for the various cancer subtypes that were analyzed in this study.

Introduction

Worldwide 2,089,000 women were diagnosed with breast cancer with an estimated related death of 626,000 in 2018.[1] Population-based breast cancer screening reduces mortality and is commonly performed with mammography.[2] However, mammography-based screening can be improved with regard sensitivity and specificity levels. It is furthermore known that tumors in dense breast tissue are often missed in a mammogram and although outweighed by mortality reduction low energy X-ray imaging carries a risk of causing radiation-induced tumors.[3] Available clinical biomarkers cancer antigen (CA) 15-3, 27-29 and 125 as well as carcinoembryonic antigen (CEA) are only of use to indicate treatment failure and are not recommended for screening, diagnosis, or staging purposes.[4] Therefore, discovery of novel biomarkers with improved test performance is widely pursued to potentially provide an add-on diagnostic tool.[5] Next to genomic markers, proteins that are present in the circulation have received great attention.[6,7] Although a large number of mass spectrometry (MS)-based exploratory studies has resulted in breast cancer protein signatures, none of these findings has been translated into a laboratory test.[8] As a consequence, biomarker strategies have been improved by properly defining the unmet clinical needs and by implementing protocols for standardized body fluid collection, high-throughput sample preparation and robust and precise MS-measurements.[5,9–12]

At the same time, MS-based proteomics studies demonstrated that post-translation modifications (PTMs) on proteins are often overlooked, although these modulate protein function and are thus an interesting source of functional biomarkers. One of the most, if not the most frequent PTMs is protein glycosylation.[13–15] Changes in protein glycosylation may have influence on or may be caused by tumor growth, differentiation, metastasis, transformation, adhesion, pathogen recognition and immune surveillance.[16,17] Protein glycosylation and its association with various cancers has been studied for more than half a century, but recent developments have allowed glyco(proteo)mics strategies to join forces with high-throughput cancer proteomics efforts to determine glycomic phenotypes and improve our understanding of the pathophysiology of various cancers.[18–24] For example, large-scale glycosylation biomarker studies based on for example immunoglobulin glycosylation and total serum *N*-glycome (TSNG) have reported changes upon cancer treatment and associations with survival.[25,26] Moreover, aberrant glycosylation profiles have been found on the surface of cancer cells with potentially diagnostic value towards evaluating tumor progression.[27,28] Breast cancer biomarker signatures have been pursued by analysis of *N*-glycan profiles in blood-derived or other body fluid samples using ultrahigh performance liquid chromatography (UPLC) methods

combined with MS identification or detection of fluorescent labels.[29–36] These studies reported associations with cancer or treatment regimes, but interestingly did not always corroborate previous findings.

In this study we report TSNG profiles from an in-house collected breast cancer cohort and compare our results with the aforementioned reports. Our sample cohort consists of 145 breast cancer patients that are age-matched with 171 healthy control individuals. *N*-glycan analysis includes linkage-specific derivatization of α2,3- and α2,6-linked sialic acids and MS-profiles are obtained using a matrix-assisted laser desorption/ionization Fourier Transform ion cyclotron resonance (MALDI-FT-ICR) platform. The potential of *N*-glycan profiles for diagnosis or staging of breast cancer is evaluated.

Materials and methods

Patients

Serum samples of 159 female patients with breast cancer and 173 female healthy volunteers were collected at the outpatients clinic at Leiden University Medical Center prior to any treatment between 2002 and 2013. The samples of the controls were matched to the cases based on age and date of sample collection. Criteria for case exclusion were; a history of cancer (other than basal cell carcinoma or cervical carcinoma in situ) shorter than 10 years before blood sampling and breast cancer in medical history. From the controls only date of birth was recorded. Table 1 provides an overview of patient characteristics and information on the invasive cancer cases (i.e. excluding ductal carcinoma in situ (DCIS). Written informed consent was obtained from patients and healthy volunteers prior to sample collection. The study was approved by the Medical Ethical Committee of the LUMC.

Serum sample collection

Blood specimens were collected and processed according to a standardized protocol. Blood was collected in a 8.5cc vacutainer serum separator tube and centrifuged for 10 min at 1000g. After centrifugation the serum was divided into 5 mL polystyrene tubes. Within 4 hours after blood collection the serum samples were stored at -80 °C. The samples underwent one freeze-thaw cycle for aliquoting in eight 60-µl tubes. All serum samples were randomly distributed in six 96-well plates, along with plasma standards (Visucon-F frozen normal control plasma, pooled from 20 human donors, citrated and buffered with 0.02 M HEPES, Affinity Biologicals, Ancaster, ON, Canada) as technical quality control samples and blanks.

Chemicals

Nonidet P-40 substitute (NP-40), potassium dihydrogenphosphate (KH_2PO_4), disodium hydrogen phosphate dihydrate ($Na_2HPO_4 \times 2H_2O$), sodium chloride (NaCl), 50% sodium hydroxide (NaOH), 1-hydroxybenzotriazole 97% (HOBt) and super-DHB (9:1 mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid, sDHB) were obtained from Sigma-Aldrich (Steinheim, Germany). Potassium hydroxide (KOH), sodium dodecyl sulfate (SDS), analytical grade ethanol and trifluoroacetic acid (TFA) were purchased from Merck (Darmstadt, Germany). HPLC-grade acetonitrile (ACN) was purchased from Biosolve (Valkenswaard, The Netherlands) and 1-ethyl-3-(3-(dimethylamino)propyl)carbodiimide (EDC) hydrochloride was obtained from Fluorochem (Hadfield, UK). Recombinant peptide-*N*-glycosidase F (PNGase F) was purchased from Roche Diagnostics (Mannheim, Germany). From a Millipore Q-Gard 2 system (Amsterdam, The Netherlands)

maintained at ≥18 MΩ milli-Q water (MQ) was generated.

Sample preparation and MALDI-FTICR-MS measurement

Enzymatic *N*-glycan release was performed as previously described.[37] In short, 6 µL sample was added to 12 µL 2% SDS and incubated for 10 min at 60 °C. After incubation 12.6 µL release mixture (6 µL 4% NP40, 6 µL 5× PBS and 0.6 µL PNGase F) was added and the samples were incubated overnight at 37 °C. The samples were stored at -20 °C before further preparation.

Ethyl esterification was performed for linkage specific stabilization of the sialic acid moieties of the glycans.[38] One microliter of released glycan sample was added to 20 µL of ethyl esterification reagent (0.25 M EDC 0.25 M HOBt in pure ethanol) and incubated for one hour at 37 °C. Subsequently 20 µL ACN was added.

Glycan purification was performed using cotton HILIC SPE microtips.[38,39] These HILIC tips were prewetted with three times 20 µL MQ and conditioned with three times 20 µL 85% ACN. Next, the sample was pipetted up and down 20 times in the HILIC tip. The HILIC phase was first washed three times with 20 µL 85% ACN with 1% TFA and second three times with 20 µL 85% ACN. Elution was performed by pipetting 10 µL MQ five times up and down. Two microliters of sample was spotted with 1 µL matrix (5 mg/mL sDHB in 50% ACN with 1 mM NaOH) onto a MALDI target (800/384 MTP AnchorChip, Bruker Daltonics, Bremen, Germany) and the spots were allowed to dry.

MALDI-FTICR-MS experiments were performed as described before.[40] A Bruker 15T solariX XR FTICR MS (Bruker Daltonics) recorded the spectra in the m/z-range from 1011.86 to 5000.00, with 1M data points. The obtained average spectra contained ten acquired scans. The system was operated by ftmsControl (version 2.1.0) software.

Data preprocessing, batch correction and statistics

Serum *N*-glycosylation profiles were obtained for 159 breast cancer patient samples and 173 healthy volunteer samples, of which respectively 145 and 171 spectra passed the quality criteria.[40] The analyte list consisted of 101 analytes which passed the quality criteria (Supporting information Table S-1). The areas of the signals were extracted using MassyTools (version 0.1.8.1). To correct for batch effects from the two MALDI-target batches (number of samples exceed the number of spots on a MALDI-target), the effect was estimated per analyte in a linear model and the values of these analytes were regressed on the MALDI-target batch (categorical variable). The standardized values were normalized to the sum of all analytes for relative quantification. Subsequently, derived traits were calculated (Supporting information Table S-2) and logistic regression was performed for each individual

glycan and each derived trait using R version 3.3.2 (R Foundation for Statistical Computing, Vienna, Austria) and RStudio, version 1.0.136 (RStudio, Boston, MA; released 21 December 2016).[41] The odds ratios (ORs) were calculated with their 95% confidence intervals (CIs) assuming a Student's t-distribution and are referring to an increase of 1 SD in the tested traits. Multivariate (principal component) analysis was performed on both individual glycans and derived traits using the various clinical parameters of the breast cancer subtypes.

Results and discussion

Serum protein *N*-glycan profiles were obtained from an in-house breast cancer cohort, consisting of 145 breast cancer cases and 171 healthy controls. In total 101 *N*-glycans were relatively quantified, including differentiation species with α2,3- and α2,6-linked sialic acids (see Materials and Methods section). Patient characteristics and information on the invasive cancer cases (i.e. excluding ductal carcinoma in situ (DCIS) is provided in Table 1. The patient group had an average age of 68 years and almost half of the group had stage II breast cancer. Quality control samples were taken along in the TSNG analysis to enable potential batch correction, as described in materials and methods.

Logistic regression analysis was performed to reveal potential differences between the glycosylation profiles of breast cancer patients and healthy controls. Moreover, it was evaluated whether glycosylation associated with one of the various clinical parameters listed in Table 1. This was done by using multivariate (principal component) analysis as well as by assuming a t-distribution of the various breast cancer subtypes. All analyses were performed for both single compositions and combined glycosylation features (further referred to as derived traits), of which the latter analysis focused on the most commonly reported cancer-associated changes in glycosylation, namely sialylation, fucosylation, and *N*-linked glycan branching.[30]

Student's t-test indicated two glycans to be lower in breast cancer patients, namely a fucosylatedtriantennary glycan that carries three α2,3-linked sialic acids (further referred to as H6N5F1L3, Figure 1, Supporting information Table S-3 and Supporting material) and a non-fucosylated triantennary glycan that carries a combination of α2,3-linked and α2,6-linked sialic acids (H6N5L2E1).Furthermore, it was found that one fucosylated tetraantennary glycan that carries a combination of α2,3-linked and α2,6-linked sialic acids (H7N6F1L1E3) was significantly elevated in breast cancer patients. Interestingly, in one previous study H6N5F1L3 has been associated with breast cancer, however in the opposite direction with elevated levels in patients as compared to controls (as is summarized in Figure 2A).[30] Similar elevated levels of triantennarytrisialylatedfucosylated glycans were reported in earlier studies, although it is emphasized that in these studies sialic acids were not determined with linkage-specificity, but rather as summarized triantennary trisialylated fucosylated glycans (referred to as H6N5F1S3, consisting of H6N5F1E3, H6N5F1L3, H6N5F1L2E1, H6N5F1L1E2 and H6N5F1E3, Supporting information Table S-3).[32,42]

In one of the older studies a significant increase was found in trisialylated triantennary glycans containing α1,3-linked fucose, pointing towards elevated levels of the sialyl-Lewis X (sLeX) epitope.[32] Similarly, Pierce and

Table 1. Patient characteristics and invasive tumor characteristics

	%	Cases (*n* = 145)	Controls (*n* = 171)
Age in years, mean (SD)		68 (13.1)	67 (11.2)
Histological type			
DCIS	16	23	n/a
Invasive ductal carcinoma	66	96	n/a
Invasive lobular carcinoma	14	21	n/a
Other	4	5	n/a
Invasive tumors (*n* = 127)			
Grade			
I	18	26	n/a
II	40	58	n/a
III	37	53	n/a
Missing	5	8	n/a
Tumor stage			
pT1	67	80	n/a
pT2	30	36	n/a
pT3/4	3	4	n/a
Nodal stage			
N0	63	77	n/a
Nmi	4	5	n/a
N1	26	32	n/a
N2	3.5	4	n/a
N3	3.5	4	n/a
Estrogen receptor (ER)- status			
Negative	16	23	n/a
Positive	68	98	n/a
Missing	16	24	n/a
Progesterone receptor (PR)-status			
Negative	54	78	n/a
Positive	30	43	n/a
Missing	16	24	n/a
Human epidermal growth factor receptor-2 (Her2)-status			
Negative	68	99	n/a
Positive	10	15	n/a
Missing	22	31	n/a

n, number of individuals; SD, standard deviation; n/a, not applicable;

co-workers reported elevated levels of agalactosylated diantennary glycans and glycans containing the sLeX epitope in patients with tumor-positive lymph nodes compared to women with no lymph node metastasis.[33] Such increased levels of the sLeX epitope in serum and on the tumorcell surface are frequently associated with cancer[30,31,43–47], but were not observed in our study upon considering different stages.

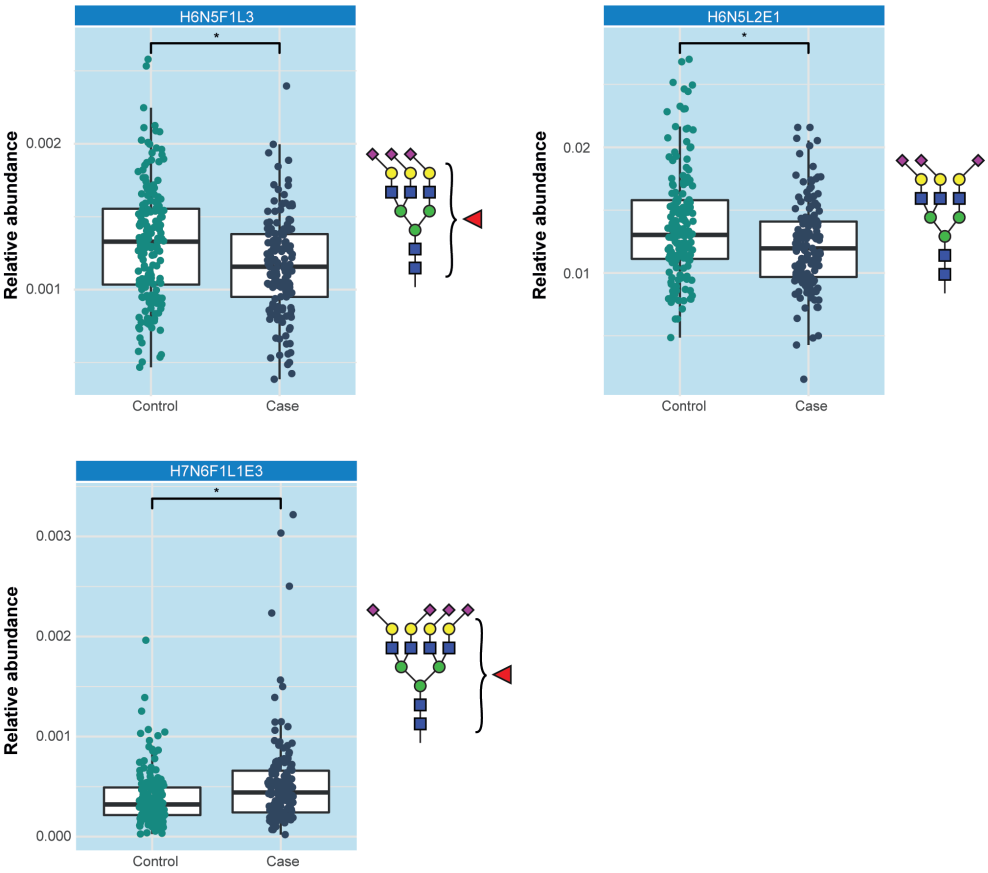


Figure 1. Association of H6N5F1L3, H6N5L2E1 and H7N6F1L1E3 with breast cancer.

With regard to the analysis of derived glycosylation traits from our data, TSNG profiles showed differences for CF, A2LF and A2F0B between breast cancer patients and healthy controls (Supporting information Table S-4). Additional differences were found when clinical parameters (Table 1) were taken into account assummarized in Figure 2B. Upon considering cancer staging, as an example the levels of oligomannose structures in breast cancer cases are plotted in Figure 3A. A trend towards a lower level of oligomannose can be seen at stage III cancer, whereas in a previous mouse study on breast cancer elevated levels of oligomannose glycans

were observed.[34] In the same study a decreased level was reported after resection and furthermorea small number of case-control human serum samples were evaluated, in which similar elevations of oligomannose glycans were observed in breast cancer patients.[34] In addition, this elevation was supported by a breast cancer cell line study .[48] Here, released glycans from cytosolic and membrane-bound glycoproteins from normal epithelial cells, invasive and non-invasive breast cancer cells were measured with MALDI-MS and the obtained profiles were compared. Notably, a decrease of oligomannose glycans in serum of breast cancer patients has also been reported[31], and literature findings on serum oligomannose glycan levels of total serum appear contradictory.

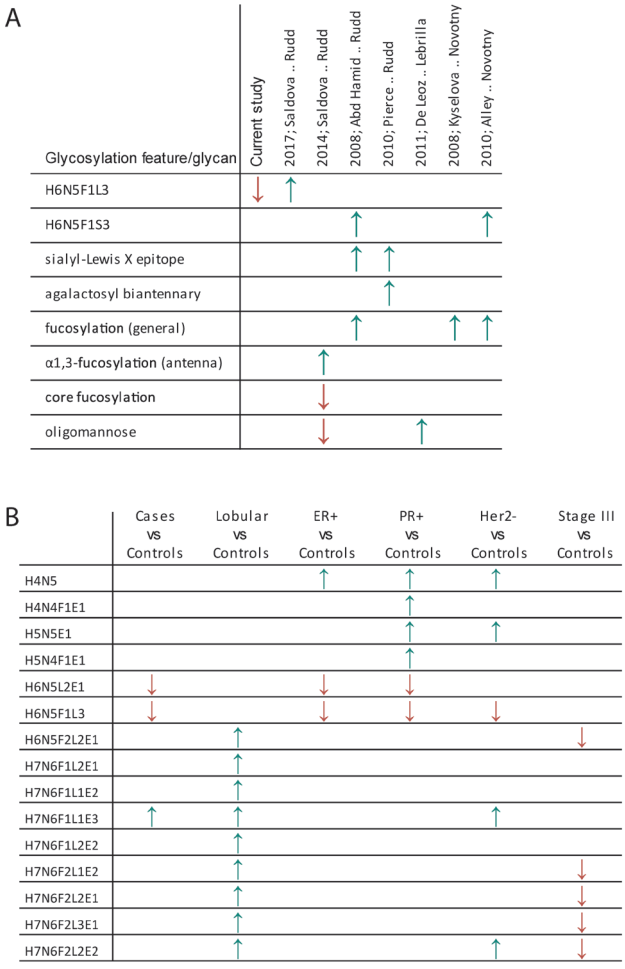


Figure 2. (A) Comparison of previously reported data and results of the current study. (B) Significant direct traits (glycan compositions) for specific breast cancer subtypes and stages as determined in a Student's t-test.

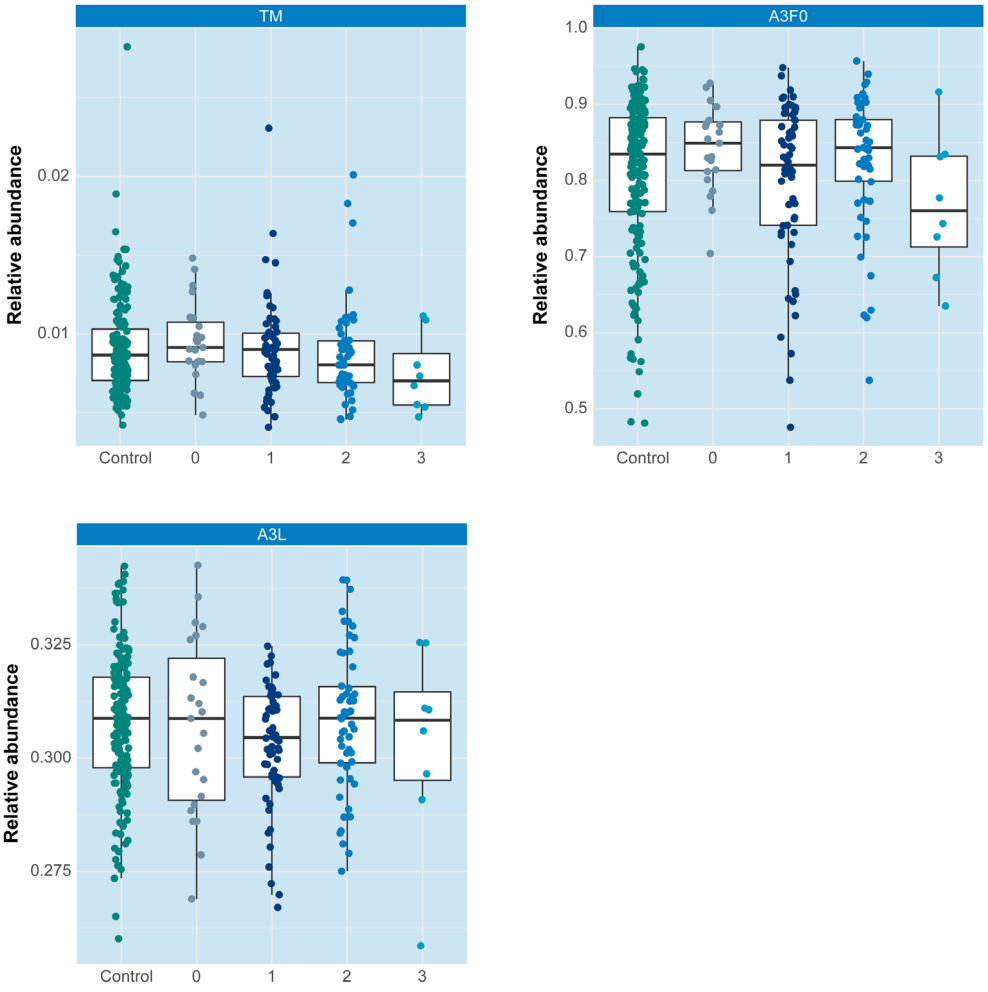


Figure 3. Three examples of derived traits and their potential association with cancer stage. Control individuals are plotted in green, patients with DCIS are plotted in grey (0), breast cancer patients are plotted in blue with staging 1=grade I, 2=grade II, 3=grade III. (A)Oligomannose structures (TM) (B) Non-fucosylated triantennary glycans (A3F0) (C)α2,3-sialylated triantennary glycans (A3L).

Results for fucosylation and sialylation traits are exemplified in Figure 3B (triantennary non-fucosylated glycans; A3F0) and Figure 3C (α2,3-sialylation of triantennary glycans;A3L), respectively. This data which is obtained from a fair number of patient samples (n=145) is not in line with previous findings of increased fucosylation and sialylation levels associated with cancer progression and staging of the disease.[29,32,42] However, when other clinical parameters are considered certain derived traits exhibit significant p-values, for example when only lobular carcinomas

are compared to controls (CF, A3F, A2LF, A3LF, A3EF and A4EF, see Supporting material). Moreover, when considering stage III patients with lobular carcinoma the levels of the three earlier mentioned glycan compositions (Figure 1) are increased by a factor of 1.5, whereas in stage III patients with ductal carcinoma these levels are decreased by a factor of 2. Although these latter observations are not significant (due to low sample numbers), this is a clear indication that the heterogeneous character of breast cancer that includes a large number of disease subtypes (as summarized in Table 1) is reflected in various *N*-glycan profiles. Of note, for our current data set, stratification according to histological subtypes did result in clear disease glycomic signatures yet. This is exemplified for fucosylation and sialylation in Figures 4A and 4B, respectively, where glycomic data are plotted separately for the two histological breast cancer types. No statistically significant were observed, possibly due to limited sample numbers. It is noted that patient cohorts in earlier studies likely consisted of different combinations of these histological subtypes. The various results reported so far emphasize the importance of detailed knowledge of clinical data and inclusion of even larger patient numbers.

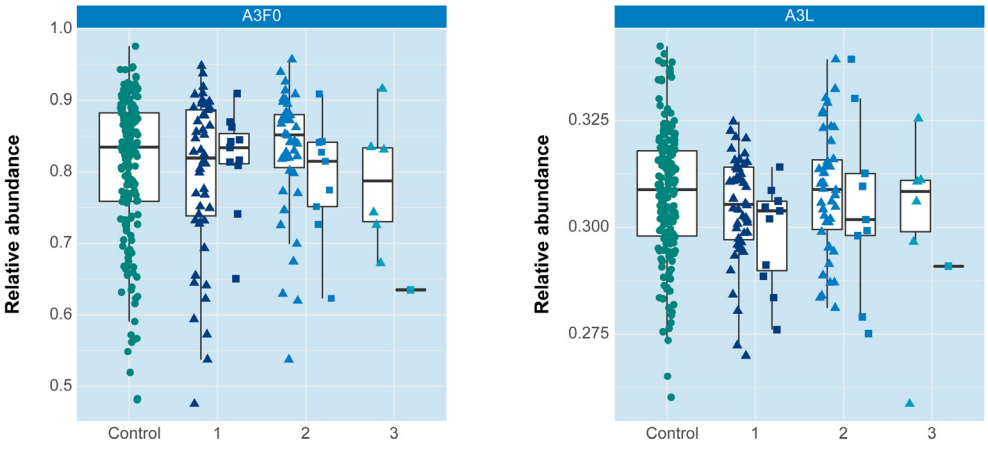


Figure 4. Two examples of derived traits and their potential association with cancer stage. Control individuals are plotted in green, breast cancer patients are plotted in blue with staging 1=grade I, 2=grade II, 3=grade III. Breast cancer patients are further stratified according to histology, namely "Invasive ductal carcinoma" (triangles) and "Invasive lobular carcinoma" (squares) (A)Non-fucosylated triantennary glycans (A3F0) (B) α2,3-sialylated triantennary glycans (A3L).

In conclusion, we have analyzed serum *N*-glycosylation profiles from breast cancer patients and healthy controls. A distinguishing signature for breast cancer was not found, although a significant difference between both groups were observed for H6N5F1L3, H6N5L2E1 and H7N6F1L1E3. In previous studies, various changes

in TSNG were reported, but also these results differed from each other and could not be replicated in our study. An evaluation of literature, together with the results of the current study, does not converge into a general breast cancer *N*-glycomic signature that distinguishes cases from controls. However, the fact that such glycomic markers are not observed can be explained by the heterogeneity of the disease and by the small size of patient cohorts. The heterogeneous character of the disease becomes clear from Table 1 that lists patients that exhibit various combinations of receptor statuses. Furthermore it is known that breast cancer tumors present a variety of histological patterns and biological characteristics.[49] In addition, the clinical response of breast cancer tumors is very different per type and up to 25% of the invasive breast cancer tumors is histologically seen a special type.[49] It is therefore recommended that in future biomarker discovery studies different subtypes within the breast cancer samples should be taken into account, instead of analyzing all breast cancer tumor subgroups together and aiming for an overarching signature.

Acknowledgements

This work was supported by the society "Genootschap ter ondersteuning van de vroege opsporing van kanker" (Lisse, The Netherlands) to further endorse the development of a blood-based test for early detection of cancer (no grant number applicable).

The authors would like to thank ing. Ronald L. van Vlierberghe (Biobank), Linda Verhoeff (Datacenter) and Elly Krol and Gemma Rankes (Outpatient clinic) from the Leiden University Medical Center for their assistance. They also would like to thank Dr. Alexia A. Kakourou and Dr. Bart J. A. Mertens from the department of Biomedical Data Sciences at the Leiden University Medical Center for performing batch correction.

References

(1) Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R. L.; Torre, L. A.; Jemal, A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA. Cancer J. Clin. 2018, 68 (6), 394–424. https://doi.org/10.3322/caac.21492.

(2) Paap, E.; Verbeek, A. L. M.; Botterweck, A. A. M.; van Doorne-Nagtegaal, H. J.; Imhof-Tas, M.; de Koning, H. J.; Otto, S. J.; de Munck, L.; van der Steen, A.; Holland, R.; den Heeten, G. J.; Broeders, M. J. M. Breast Cancer Screening Halves the Risk of Breast Cancer Death: A Case-Referent Study. Breast 2014, 23 (4), 439–444. https://doi.org/10.1016/j.breast.2014.03.002.

(3) Warren, L. M.; Dance, D. R.; Young, K. C. Radiation Risk of Breast Screening in England with Digital Mammography. Br. J. Radiol. 2016, 89 (1067). https://doi.org/10.1259/bjr.20150897.

(4) Harris, L.; Fritsche, H.; Mennel, R.; Norton, L.; Ravdin, P.; Taube, S.; Somerfield, M. R.; Hayes, D. F.; Bast, R. C. American Society of Clinical Oncology 2007 Update of Recommendations for the Use of Tumor Markers in Breast Cancer. J. Clin. Oncol. 2007, 25 (33), 5287–5312. https://doi.org/10.1200/JCO.2007.14.2364.

(5) Lord, S. J.; St John, A.; Bossuyt, P. M. M.; Sandberg, S.; Monaghan, P. J.; O'Kane, M.; Cobbaert, C. M.; Röddiger, R.; Lennartz, L.; Gelfi, C.; Horvath, A. R.; Test Evaluation Working Group of the European Federation of Clinical Chemistry and Laboratory Medicine. Setting Clinical Performance Specifications to Develop and Evaluate Biomarkers for Clinical Use. Ann. Clin. Biochem. 2019, 56 (5), 527–535. https://doi.org/10.1177/0004563219842265.

(6) Rontogianni, S.; Synadaki, E.; Li, B.; Liefaard, M. C.; Lips, E. H.; Wesseling, J.; Wu, W.; Altelaar, M. Proteomic Profiling of Extracellular Vesicles Allows for Human Breast Cancer Subtyping. Commun. Biol. 2019, 2 (1), 325. https://doi.org/10.1038/s42003-019-0570-8.

(7) Anderson, L. Within Sight of a Rational Pipeline for Development of Protein Diagnostics. Clin. Chem. 2012, 58 (1), 28–30. https://doi.org/10.1373/clinchem.2011.173377.

(8) Lam, S. W.; Jimenez, C. R.; Boven, E. Breast Cancer Classification by Proteomic Technologies: Current State of Knowledge. Cancer Treat. Rev. 2014, 40 (1), 129–138. https://doi.org/10.1016/j.ctrv.2013.06.006.

(9) Chen, X.; Flynn, G. C. Analysis of N-Glycans from Recombinant Immunoglobulin G by on-Line Reversed-Phase High-Performance Liquid Chromatography/Mass Spectrometry. Anal. Biochem. 2007, 370 (2), 147–161. https://doi.org/10.1016/j.ab.2007.08.012.

(10) Belczacka, I.; Latosinska, A.; Metzger, J.; Marx, D.; Vlahou, A.; Mischak, H.; Frantzi, M. Proteomics Biomarkers for Solid Tumors: Current Status and Future Prospects. Mass Spectrom. Rev. 2019, 38 (1), 49–78. https://doi.org/10.1002/mas.21572.

(11) Abbatiello, S. E.; Schilling, B.; Mani, D. R.; Zimmerman, L. J.; Hall, S. C.; MacLean, B.; Albertolle, M.; Allen, S.; Burgess, M.; Cusack, M. P.; Gosh, M.; Hedrick, V.; Held, J. M.; Inerowicz, H. D.; Jackson, A.; Keshishian, H.; Kinsinger, C. R.; Lyssand, J.; Makowski, L.; Mesri, M.; Rodriguez, H.; Rudnick, P.; Sadowski, P.; Sedransk, N.; Shaddox, K.; Skates, S. J.; Kuhn, E.; Smith, D.; Whiteaker, J. R.; Whitwell, C.; Zhang, S.; Borchers, C. H.; Fisher, S. J.; Gibson, B. W.; Liebler, D. C.; MacCoss, M. J.; Neubert, T. A.; Paulovich, A. G.; Regnier, F. E.; Tempst, P.; Carr, S. A. Large-Scale Interlaboratory Study to Develop, Analytically Validate and Apply Highly Multiplexed, Quantitative Peptide Assays to Measure Cancer-Relevant Proteins in Plasma. Mol. Cell. Proteomics 2015, 14 (9), 2357–2374. https://doi.org/10.1074/mcp.M114.047050.

(12) Wang, H.; Shi, T.; Qian, W. J.; Liu, T.; Kagan, J.; Srivastava, S.; Smith, R. D.; Rodland, K. D.; Camp, D. G. The Clinical Impact of Recent Advances in LC-MS for Cancer Biomarker Discovery and Verification. Expert Rev. Proteomics 2016, 13 (1), 99–114. https://doi.org/10.1586/14789450.2016.1122529.

(13) Varki, A.; Gagneux, P. Biological Functions of Glycans. In Essentials of Glycobiology; Varki, A., Ed.; Cold Spring Harbor: New York, 2017.

(14) Smith, L. M.; Kelleher, N. L. Proteoforms as the next Proteomics Currency. Science (80-.). 2018, 359 (6380), 1106–1107. https://doi.org/10.1126/science.aat1884.

(15) Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; Ge, Y.; Gunawardena, J.; Hendrickson, R. C.; Hergenrother, P. J.; Huber, C. G.; Ivanov, A. R.; Jensen, O. N.; Jewett, M. C.; Kelleher, N. L.; Kiessling, L. L.; Krogan, N. J.; Larsen, M. R.; Loo, J. A.; Ogorzalek Loo, R. R.; Lundberg, E.; Maccoss, M. J.; Mallick, P.; Mootha, V. K.; Mrksich, M.; Muir, T. W.; Patrie, S. M.; Pesavento, J. J.; Pitteri, S. J.; Rodriguez, H.; Saghatelian, A.; Sandoval, W.; Schlüter, H.; Sechi, S.; Slavoff, S. A.; Smith, L. M.; Snyder, M. P.; Thomas, P. M.; Uhlén, M.; Van Eyk, J. E.; Vidal, M.; Walt, D. R.; White, F. M.; Williams, E. R.; Wohlschlager, T.; Wysocki, V. H.; Yates, N. A.; Young, N. L.; Zhang, B. How Many Human Proteoforms Are There? Nat. Chem. Biol. 2018, 14 (3), 206–214. https://doi.org/10.1038/nchembio.2576.

(16) Almeida, A.; Kolarich, D. The Promise of Protein Glycosylation for Personalised Medicine. Biochim. Biophys. Acta - Gen. Subj. 2016, 1860 (8), 1583–1595. https://doi.org/10.1016/j.bbagen.2016.03.012.

(17) Dotz, V.; Wuhrer, M. N -glycome Signatures in Human Plasma: Associations with Physiology and Major Diseases . FEBS Lett. 2019, 1–11. https://doi.org/10.1002/1873-3468.13598.

(18) Knežević, A.; Polašek, O.; Gornik, O.; Rudan, I.; Campbell, H.; Hayward, C.; Wright, A.; Kolčić, I.; O'Donoghue, N.; Bones, J.; Rudd, P. M.; Lauc, G. Variability, Heritability and Environmental Determinants of Human Plasma N-Glycome. J. Proteome Res. 2009, 8 (2), 694–701. https://doi.org/10.1021/pr800737u.

(19) Drake, P. M.; Cho, W.; Li, B.; Prakobphol, A.; Johansen, E.; Anderson, N. L.; Regnier, F. E.; Gibson, B. W.; Fisher, S. J. Sweetening the Pot: Adding Glycosylation to the Biomarker Discovery Equation. Clin. Chem. 2010, 56 (2), 223–236. https://doi.org/10.1373/clinchem.2009.136333.

(20) Kirwan, A.; Utratna, M.; O'Dwyer, M. E.; Joshi, L.; Kilcoyne, M. Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. Biomed Res. Int. 2015, 2015, 490531. https://doi.org/10.1155/2015/490531.

(21) Kailemia, M. J.; Park, D.; Lebrilla, C. B. Glycans and Glycoproteins as Specific Biomarkers for Cancer. Anal. Bioanal. Chem. 2017, 409 (2), 395–410. https://doi.org/10.1007/s00216-016-9880-6.

(22) Kailemia, M. J.; Xu, G.; Wong, M.; Li, Q.; Goonatilleke, E.; Leon, F.; Lebrilla, C. B. Recent Advances in the Mass Spectrometry Methods for Glycomics and Cancer. Anal. Chem. 2018, 90 (1), 208–224. https://doi.org/10.1021/acs.analchem.7b04202.

(23) Peng, W.; Zhao, J.; Dong, X.; Banazadeh, A.; Huang, Y.; Hussien, A.; Mechref, Y. Clinical Application of Quantitative Glycomics. Expert Rev. Proteomics 2018, 15 (12), 1007–1031. https://doi.org/10.1080/14789450.2018.1543594.

(24) Song, E.; Mechref, Y. Defining Glycoprotein Cancer Biomarkers by MS in Conjunction with Glycoprotein Enrichment. Biomark. Med. 2015, 9 (9), 835–844. https://doi.org/10.2217/bmm.15.55.

(25) Breen, L. D.; Pučić-Baković, M.; Vučković, F.; Reiding, K.; Trbojević-Akmačić, I.; Gajdošik, M. Š.; Cook, M. I.; Lopez, M. J.; Wuhrer, M.; Camara, L. M.; Andjelković, U.; Dupuy, D. E.; Josić, D. IgG and IgM Glycosylation Patterns in Patients Undergoing Image-Guided Tumor Ablation. Biochim. Biophys. Acta - Gen. Subj. 2016, 1860 (8), 1786–1794. https://doi.org/10.1016/j.bbagen.2016.01.011.

(26) de Vroome, S. W.; Holst, S.; Girondo, M. R.; van der Burgt, Y. E. M.; Mesker, W. E.; Tollenaar, R. A. E. M.; Wuhrer, M. Serum N-Glycome Alterations in Colorectal Cancer Associate with Survival. Oncotarget 2018, 9 (55), 30610–30623. https://doi.org/10.18632/oncotarget.25753.

(27) Park, H.; Hwang, M. P.; Kim, Y.; Kim, K. J.; Jin, J. M.; Kim, Y. H.; Yang, Y. H.; Lee, K. H.; Kim, Y. G. Mass Spectrometry-Based N-Linked Glycomic Profiling as a Means for Tracking Pancreatic Cancer Metastasis. Carbohydr. Res. 2015, 413 (1), 5–11. https://doi.org/10.1016/j.carres.2015.04.019.

(28) Lauc, G.; Wuhrer, M. High-Throughput Glycomics and Glycoproteomics; Lauc, G., Wuhrer, M., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2017; Vol. 1503. https://doi.org/10.1007/978-1-4939-6493-2.

(29) Kyselova, Z.; Mechref, Y.; Kang, P.; Goetz, J. A.; Dobrolecki, L. E.; Sledge, G. W.; Schnaper, L.; Hickey, R. J.; Malkas, L. H.; Novotny, M. V. Breast Cancer Diagnosis and Prognosis through Quantitative Measurements of Serum Glycan Profiles. Clin. Chem. 2008, 54 (7), 1166–1175. https://doi.org/10.1373/clinchem.2007.087148.

(30) Saldova, R.; Haakensen, V. D.; Rødland, E.; Walsh, I.; Stöckmann, H.; Engebraaten, O.; Børresen-Dale, A. L.; Rudd, P. M. Serum N-Glycome Alterations in Breast Cancer during Multimodal Treatment and Follow-Up. Mol. Oncol. 2017, 11 (10), 1361–1379. https://doi.org/10.1002/1878-0261.12105.

(31) Saldova, R.; Asadi Shehni, A.; Haakensen, V. D.; Steinfeld, I.; Hilliard, M.; Kifer, I.; Helland, Å.; Yakhini, Z.; Børresen-Dale, A. L.; Rudd, P. M. Association of N-Glycosylation with Breast Carcinoma and Systemic Features Using High-Resolution Quantitative UPLC. J. Proteome Res. 2014, 13 (5), 2314–2327. https://doi.org/10.1021/pr401092y.

(32) Abd Hamid, U. M.; Royle, L.; Saldova, R.; Radcliffe, C. M.; Harvey, D. J.; Storr, S. J.; Pardo, M.; Antrobus, R.; Chapman, C. J.; Zitzmann, N.; Robertson, J. F.; Dwek, R. A.; Rudd, P. M. A Strategy to Reveal Potential Glycan Markers from Serum Glycoproteins Associated with Breast Cancer Progression. Glycobiology 2008, 18 (12), 1105–1118. https://doi.org/10.1093/glycob/cwn095.

(33) Pierce, A.; Saldova, R.; Abd Hamid, U. M.; Abrahams, J. L.; McDermott, E. W.; Evoy, D.; Duffy, M. J.; Rudd, P. M. Levels of Specific Glycans Significantly Distinguish Lymph Node-Positive from Lymph Node-Negative Breast Cancer Patients. Glycobiology 2010, 20 (10), 1283–1288. https://doi.org/10.1093/glycob/cwq090.

(34) de Leoz, M. L. A.; Young, L. J. T. T.; An, H. J.; Kronewitter, S. R.; Kim, J.; Miyamoto, S.; Borowsky, A. D.; Chew, H. K.; Lebrilla, C. B. High-Mannose Glycans Are Elevated during Breast Cancer Progression. Mol. Cell. Proteomics 2011, 10 (1), 1–9. https://doi.org/10.1074/mcp.M110.002717-1.

(35) Terkelsen, T.; Haakensen, V. D. V. D.; Saldova, R.; Gromov, P.; Hansen, M. K. M. K.; Stöckmann, H.; Lingjærde, O. C. O. C.; Børresen-Dale, A. L. A.-L. A. L.; Papaleo, E.; Helland, Å.; Rudd, P. M. P. M.; Gromova, I.; Helland,

A.; Rudd, P. M. P. M.; Gromova, I.; Helland, Å.; Rudd, P. M. P. M.; Gromova, I. N-Glycan Signatures Identified in Tumor Interstitial Fluid and Serum of Breast Cancer Patients: Association with Tumor Biology and Clinical Outcome. Mol. Oncol. 2018, 12 (6), 972–990. https://doi.org/10.1002/1878-0261.12312.

(36) Peng, W.; Goli, M.; Mirzaei, P.; Mechref, Y. Revealing the Biological Attributes of N-Glycan Isomers in Breast Cancer Brain Metastasis Using Porous Graphitic Carbon (PGC) Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS). J. Proteome Res. 2019, 18 (10), 3731–3740. https://doi.org/10.1021/acs.jproteome.9b00429.

(37) Bladergroen, M. R.; Reiding, K. R.; Hipgrave Ederveen, A. L.; Vreeker, G. C. M.; Clerc, F.; Holst, S.; Bondt, A.; Wuhrer, M.; van der Burgt, Y. E. M. Automation of High-Throughput Mass Spectrometry-Based Plasma N -Glycome Analysis with Linkage-Specific Sialic Acid Esterification. J. Proteome Res. 2015, 14 (9), 4080–4086. https://doi.org/10.1021/acs.jproteome.5b00538.

(38) Reiding, K. R.; Blank, D.; Kuijper, D. M.; Deelder, A. M.; Wuhrer, M. High-Throughput Profiling of Protein N-Glycosylation by MALDI-TOF-MS Employing Linkage-Specific Sialic Acid Esterification. Anal. Chem. 2014, 86 (12), 5784–5793. https://doi.org/10.1021/ac500335t.

(39) Selman, M. H. J.; Hemayatkar, M.; Deelder, A. M. A. M.; Wuhrer, M. Cotton HILIC SPE Microtips for Microscale Purification and Enrichment of Glycans and Glycopeptides. Anal. Chem. 2011, 83 (7), 2492–2499. https://doi.org/10.1021/ac1027116.

(40) Vreeker, G. C. M.; Nicolardi, S.; Bladergroen, M. R.; van der Plas, C. J.; Mesker, W. E.; Tollenaar, R. A. E. M.; van der Burgt, Y. E. M.; Wuhrer, M. Automated Plasma Glycomics with Linkage-Specific Sialic Acid Esterification and Ultrahigh Resolution MS. Anal. Chem. 2018, 90 (20), 11955–11961. https://doi.org/10.1021/acs.analchem.8b02391.

(41) Chong, J.; Xia, J. MetaboAnalystR: An R Package for Flexible and Reproducible Analysis of Metabolomics Data. Bioinformatics 2018, 34 (24), 4313–4314. https://doi.org/10.1093/bioinformatics/bty528.

(42) Alley, W. R.; Madera, M.; Mechref, Y.; Novotny, M. V. Chip-Based Reversed-Phase Liquid Chromatography-Mass Spectrometry of Permethylated N-Linked Glycans: A Potential Methodology for Cancer-Biomarker Discovery. Anal. Chem. 2010, 82 (12), 5095–5106. https://doi.org/10.1021/ac100131e.

(43) Saldova, R.; Reuben, J. M.; Abd Hamid, U. M.; Rudd, P. M.; Cristofanilli, M. Levels of Specific Serum N-Glycans Identify Breast Cancer Patients with Higher Circulating Tumor Cell Counts. Ann. Oncol. 2011, 22 (5), 1113–1119. https://doi.org/10.1093/annonc/mdq570.

(44) Kossowska, B.; Ferens-Sieczkowska, M.; Gancarz, R.; Passowicz-Muszyńska, E.; Jankowska, R. Fucosylation of Serum Glycoproteins in Lung Cancer Patients. Clin. Chem. Lab. Med. 2005, 43 (4), 361–369. https://doi.org/10.1515/CCLM.2005.066.

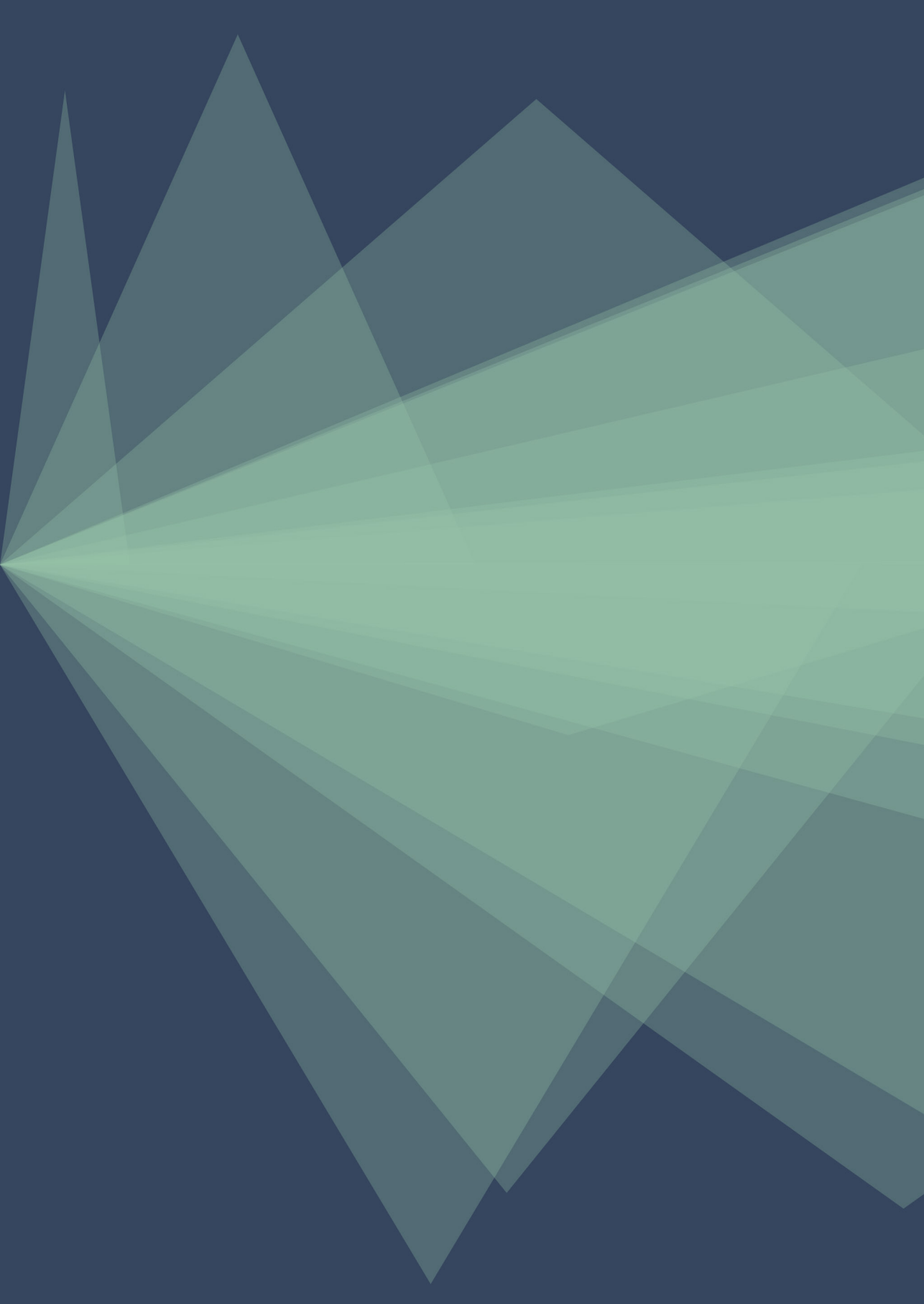
(45) Reis, C. A.; Osorio, H.; Silva, L.; Gomes, C.; David, L. Alterations in Glycosylation as Biomarkers for Cancer Detection. J. Clin. Pathol. 2010, 63 (4), 322–329. https://doi.org/10.1136/jcp.2009.071035.

(46) Bassagañas, S.; Pérez-Garay, M.; Peracaula, R. Cell Surface Sialic Acid Modulates Extracellular Matrix Adhesion and Migration in Pancreatic Adenocarcinoma Cells. Pancreas 2014, 43 (1), 109–117. https://doi.org/10.1097/MPA.0b013e31829d9090.

(47) Arnold, J. N.; Saldova, R.; Abd Hamid, U. M.; Rudd, P. M. Evaluation of the Serum N-Linked Glycome for the Diagnosis of Cancer and Chronic Inflammation. Proteomics 2008, 8 (16), 3284–3293. https://doi.org/10.1002/pmic.200800163.

(48) Goetz, J. A.; Mechref, Y.; Kang, P.; Jeng, M. H.; Novotny, M. V. Glycomic Profiling of Invasive and Non-Invasive Breast Cancer Cells. Glycoconj. J. 2009, 26 (2), 117–131. https://doi.org/10.1007/s10719-008-9170-4.

(49) Weigelt, B.; Reis-Filho, J. S. Histological and Molecular Types of Breast Cancer: Is There a Unifying Taxonomy? Nat. Rev. Clin. Oncol. 2009, 6 (12), 718–730. https://doi.org/10.1038/nrclinonc.2009.166.



CHAPTER 7

Discussion and future perspectives

From method development to a standardized protocol

In this thesis, automated analysis of released *N*-glycans from serum and plasma proteins is described (Chapter 2). This analytical method was extended with a protocol for the analysis of released glycans from dried blood spots (DBS) (Chapter 3) and the MS-part of the protocol was further elaborated on, with increased resolving power of the measurements that elucidated overlapping signals of isobaric *N*-glycans and small *O*-glycans in the MALDI-matrix region of the mass spectrum (Chapter 4).

Sample preparation, automation and MS measurements

Method development involves a stepwise process in which an existing protocol is further optimized or alternatively a method is established from scratch. Both approaches require a combination of new knowledge with the latest innovations. For the automated *N*-glycosylation profiling method that was developed in this study, a previously established protocol was used as a starting point.[1] The initial protocol was improved through multiple steps. To begin with, the pH in the enzymatic release reaction was lowered to prevent side reactions in the subsequent derivatization reaction. Second, a manual solid phase extraction (SPE) protocol based on cotton hydrophilic interaction liquid chromatography (HILIC) that is generally used in our research group[2,3] was converted into an automated version. And last, MALDI-FTICR-MS was introduced for glycan profiling instead of MALDI-TOF-MS to improve resolving power and accuracies of the mass measurements.

In a previous study multiple byproducts and fragmentations of glycans were reported after detailed analysis of the glycan mass spectra, which complicate the interpretation (peak assignments) of the spectrum.[4] Although application of MALDI-FTICR-MS allowed resolving overlapping signals, still additional signals resulting from unwanted side-reactions were present in the spectra that potentially bias the (relative) quantitation part. To prevent side reactions at the reducing end of the glycan, the pH of the release solution was lowered to enhance hydrolysis of the enzymatic release product, which contains an amine group at the reducing end (Chapter 2). This had to be carefully balanced, because a low pH can lead to protein precipitation, including the precipitation of the release enzyme. Moreover, the activity of protein *N*-glycosidase F (PNGase F) can be affected by a lower pH, or turn inactive at pH<4.0.[5]

The inclusion and automation of a cotton HILIC-based purification instead of the previous filter-plate method was important for the glycan detectability: from experience it was known that in MALDI-TOF-MS analysis the glycans in the *m*/*z* range larger than 3100 were lower in intensity when purified with the previous

method[1] compared to manual cotton HILIC methods.[2,3] This difference can be explained by either a lower recovery of larger glycans or a suppressing effect from other compounds present after filter-plate purification. The automation of a manual protocol is often not a one-to-one transferable process and in the case of cotton HILIC purification adjustments were needed. In manual pipetting processes of the samples the liquid level and SPE material can be carefully controlled, whereas this is standardized in an automated procedure. Especially for cotton HILIC SPE this was challenging since the cotton strand should remain at the end of the tip at all time during the pipetting procedure. Manipulation of the liquid can loosen the cotton strand and visually checking the potential movement of this strand is necessary. To solve this for the automated pipetting platform, porous frits were added to the pipetting tips, forcing the cotton strand to stay at the end of the tip, while letting the liquid pass. Control of the purification was ensured without the need of human involvement. This change of purification method resulted in improved detection of glycans in the higher mass range (m/z>3100), which especially in combination with MALDI-FTICR-MS resulted into the inclusion of more *N*-glycan species in relative quantification analysis.

Glycosylation profiling with MALDI-FTICR-MS allowed the detection of 112 different glycan compositions, with MALDI-TOF-MS at 86 different compositions (Chapter 2).[6] Although the FTICR-MS spectra exhibit a high resolving power, not all signals were fully resolved with the initial measurement method. Importantly, these glycan signals could be identified upon implementation of absorption mode data processing instead of the initially used magnitude mode (Chapter 4).[7] It is noted that absorption mode spectra improve identifications in profiling, but are not suited for relative quantification due to "negative" peaks near the analyte signals that originate from phasing artefacts. In future, absorption mode measurements might be used for relative quantification, when phasing software is able to prevent this baseline distortion or extraction software is able to cope with these peaks. Still, for the detection and confident identification of overlapping species this measurement mode is a valuable addition to the magnitude mode measurements.

Sample collection and the importance of standardization

The developed released *N*-glycan protocol can be widely applied to any serum or plasma cohort for which *N*-glycosylation profiles are the anticipated data (Chapter 2). This was exemplified in a pancreatic cancer case-control cohort (Chapter 5). In addition, the released *N*-glycan protocol is suited for DBS samples with a specific extension (Chapter 3), which would be an easier way of collecting samples from patients, as it even can be collected by the patient self.

Most biomarker studies have been carried out on retrospective cohorts

of clinical blood samples, that is material previously collected. The pre-analytical variables are therefore not necessarily optimal for the analytes and/or research question. Moreover, the primary concern of the clinician lies on the health and well-being of the patient and as a consequence the procedure and execution of sample collection has lower priority.[8] Research efforts and anticipated clinical diagnostics increasingly shift from a wide range of generally applied biomarkers to specific biomarkers analyzed in high-end assays to aim for personalized medicine and study the biology of an individual. As the outcomes of these type of assays may be affected by the quality of the collected sample, analytical scientists should bring their quality requirements into the clinical field.[8]

A critical aspect of sample collection is that it requires standardization both when a patient sample is acquired in the clinic, but also in case the patient performs self-sampling. Without a clear and unambiguous protocol for sample collection and -storage, variation may be introduced, resulting in biases or random variation in the results. The standardization of the pre-analytical procedures is therefore as important as the standardization of the analytical method itself. The samples in the present studies (Chapter 5 and 6) were collected according to a standardized protocol and the pre-analytical procedures were performed in a specialized laboratory. However, not every sample collection center has this type of facility and resources, but also in a standard lab facility high levels of standardization should be urged. Moreover, there should be aimed for a national, European or even global type of standardization in sample collection for different types of analytes. To get this done, it might take a long time and much effort, but the great advantage of this would be the ability to collect large sample cohorts, also for relatively rare diseases, which will improve biomarker research.

In the standardization of a sampling procedure, it should be considered that patient lifestyle and also care decisions influence the sample and that the analytical profiles obtained from a patient may not represent the actual health status of a patient.[8] Additionally, when blood is in a tube, biological (enzymes), chemical (hydrolysis, oxidation) and physical (aggregation) degradation can occur in the sample, which might lead to alterations of the biomarkers in the sample.[8,9] However, MALDI-TOF-MS glycosylation profiles have shown robustness with regard to fluctuations in time before centrifugation of blood specimens, serum and plasma storage times, storage temperatures and tube additives (such as EDTA).[10]

Besides the sampling protocol, other aspects need to be considered, such as the type of sample vials used and the storage conditions. The type of matrix in the sample container or the material of the container itself can influence the analytes and therefore the outcome of the analysis.[8] Additionally, the storage temperature can influence the degradation of the sample. It is, for example, known that in case of

protein analysis storage temperatures have significant effects on protein quantities.[11] Notably, glycosylation profiles did not differ for different storage temperatures. In addition, for DBS it was shown that even storage at 37 °C did not alter the glycosylation profile (Chapter 3).

To implement DBS collection into the clinic for patient sample collection spot quality needs to be further investigated. As the samples will be collected by different people (preferably the patients) information on 'good' and 'bad' spot quality for glycosylation analysis needs to be available. In addition, the stability of glycosylation profiles over a longer time period should be studied, as spots might need to be preserved as reference sample for new measurements in case of longitudinal sample collection.

Longitudinal samples can be measured at various time points without inclusion of the previous sample as a reference, provided the performance of the analytical protocol and instruments is kept stable over time. The technical variability over time should be minimized and a short-term version of this was already shown (Chapter 2). The technical variation between sample-batches and measurement days was determined, resulting in a less than 10% average variation of the ten most abundant glycans in the spectra.[6] We see this as an acceptable variation in our method, however lowering this number would make our method more sensitive for biological variation, which is in cohort analysis of course the preferred variation found. With the current technical variation probably subtle differences between cases and controls are missed, while these could provide essential information about glycosylation changes in disease. To decrease technical variation, attention should be paid to the last part of the protocol: MALDI-spotting and MALDI-MS analysis. It is known that sDHB-matrix does not dry as a homogeneous crystal layer, but very inhomogeneous with large crystals and partially empty spots. Averaging multiple shots at the spots helps to obtain a relatively stable spectrum, but still variation is introduced. The development of a homogeneously drying matrix suitable for released glycan measurement might therefore be necessary to decrease the total technical variation of measurements in the future.

The use of software for data-processing and interpretation

Multiple software tools were used for high-throughput data processing and -analysis of the obtained MS spectra. One of the software tools used throughout this thesis is in-house developed MassyTools (Chapter 2, 3, 5 and 6).[12] Additionally, specific scripts in R software tool were available for the calculation of derived traits, which enables us to analyze shared features of glycans, such as the branching, fucosylation, sialylation and bisection (Chapter 5 and 6).

These tools are useful and essential for fast processing of mass spectrometry

data from for example released glycans. In addition, the automated workflow minimizes the human intervention, for which also the risk of human mistakes is decreased. However, often human involvement is required for interpretation and to monitor the critical steps in a procedure. The software can still output processed data, without necessarily showing that there was an irregularity in the processing. This was for example the case in our MassyTools sample processing, where peak extraction is performed directly after calibration. An overview of the quality of the calibration of all spectra is not displayed (it is outputted per spectrum), although this is information should be evaluated before the start of the extraction. A bad calibration can highly influence the data obtained after extraction and therefore also the results of a study. For this reason, in the described studies the parts-per-million (PPM)-errors of the calibrants in all spectra were evaluated manually before peak extraction. The summarized output of crucial information about the quality of each processing step would be a valuable and important addition to data processing software tools.

Understanding of the software and the acquired data is essential for the right interpretation of the data and to answer the research questions with the right information. Therefore, software should not be treated as a black box, but be evaluated if it meets the purpose of the research and fits the type of data. To illustrate this, software can be used to find an elemental composition for an observed m/z[13], however, these type of software tools often do not differentiate between molecular compositions that are possible according to chemistry and random compositions coincidently matching the m/z.

Also, in glycosylation and MS analysis software tools should be used with caution as a tool to help human work but not to replace it completely. In glycosylation analysis the biosynthetic pathway should be taken into account when identifying MS-signals, as the inclusion of compositions that are not in accordance with the biosynthesis can complicate the interpretation of the data or spread 'fake' results into the field. Additionally, user-defined settings should be re-evaluated when changing systems. An example for this is the extraction width (in the mass window) of signals in MALDI-TOF-MS and MALDI-FTICR-MS spectra. For MALDI-TOF-MS spectra a single extraction width was sufficient, whereas for MALDI-FTICR-MS every analyte required a specific extraction window (Chapter 2).

A certain understanding of glycobiology is important for interpretation of results, as is demonstrated in this thesis (Chapter 5). For example, in the pancreatic cancer study, glycan traits A3F0S and A4F0S showed a decrease in patients while A3FS and A4FS were observed at elevated levels. When interpreting each trait individually it could be concluded that sialylation levels are changing in patients. However, by interpretation of all traits together it is far more likely that the effect

is caused by the change in fucosylation. In the case of A4F0E this might also be the case, as generally the α2,6-linked sialylation in increasing, but the non-fucosylated fraction of glycans is decreasing and appears to be the stronger effect in these glycans. Additionally, derived traits only make sense when most of the glycans in one trait show the same effect, and the resulting change described in this derived trait is not the result of for example one single glycan. Without human interpretation of the data the conclusions on this data could have been significantly different.

Limitations of total plasma *N*-glycome analysis

MALDI-FTICR-MS analysis of released *N*-glycans is a useful method for screening changes in the *N*-glycome of blood-based samples. The method is rapid, almost fully automated and exhibits a good repeatability. However, as every method, this method also has its limitations.

The most important one is that the glycan information represents a summary of all glycoproteins in the blood. The released glycans are obtained from many different glycoproteins and end up in the same mixture, so the glycans cannot be traced back to their originating glycoprotein or to the glycosylation site on that protein. Although for some glycoproteins it is known what type of glycans they carry[14], it can only be speculated from a total serum *N*-glycome profile which glycoproteins are involved in causing specific signatures and glycosylation differences. Nonetheless, global serum glycosylation analysis can still indicate what the potentially interesting glycoproteins for a follow-up study would be and give insights in major changes occurring in disease.

A second limitation is found in the MALDI-MS readout that cannot distinguish all potential isomers, although sialic acids are detected with linkage-specificity, due to the derivatization method applied. For other isomers complementary methods are required, for example separation methods (e.g. liquid chromatography or capillary electrophoresis). In cancer analysis (Chapter 5 and 6), it would for example be beneficial to know the location of the fucose on a glycan, as it turns out that fucosylation is strongly increased in pancreatic cancer (Chapter 5), but that from the current data it cannot be derived whether this increase solely involves antenna-fucoses and may be linked to the formation of sialyl-Lewis X (sLeX) moieties or that also core fucosylation is changing.

Clinical translation

Many studies on biomarker research have been performed in the past and although a great potential was foreseen, only a very limited number of biomarkers have successfully been translated into the clinic.[15,16] Multiple reasons for this translation gap have been discussed, such as the complexity of diseases and lack of knowledge about the specific diseases. In addition, technical challenges and the use of inappropriate samples and study design for the research question or validation have been mentioned.[17] The latter is also related to the collection of the samples, which has been discussed in this thesis. Here, close collaboration between clinicians, the clinical laboratory and research scientists is important.

Translational research

A screening test should meet specific requirements, of which one of the most important is that it should identify a disease when it is present (sensitivity) and rule out the disease when it is not present (specificity). In many cancer biomarker studies this criterion is not met sufficiently, leading to false positives and false negatives.[18] To bring a test into the clinic, this clinical validity is highly important as it should perform better than the currently used diagnostic method, which is often called the 'gold standard'.[19] A new clinical test should also generally provide benefits as compared to the previous one. For example, it should not include disproportionate increase of invasiveness or risk for patients. In addition, the analytical validity of a test, as was already discussed above, is highly important for a method to be transferred into the clinic.

As mentioned earlier, a limited number of biomarkers were successfully translated into a clinical test and multiple reasons for this where mentioned.[15,17] For this reason, an international initiative suggested that for the diagnosis of diseases improvements in biomarker selection should be made and that clinical development should focus more on unmet clinical needs.[20] The focus of biomarker research is not always directed towards shortcomings in clinical practice, but might also be influenced by non-clinical factors such as technological innovation (not necessarily for healthcare purposes) and financial pressure.[20] To stimulate the translation of biomarkers into the clinic in future, efforts should be made to deal with these aspects of research.

Cancer biology and blood-based cancer detection

Tumors are complex tissues, consisting of multiple cell types with various types of interactions.[21] Not only the tumor cells themselves, but also the surrounding tissues are involved in the development of the cancer.[21] Studies of human cancers

have shown that development of tumors is a result of subsequent genetic changes which provide an advantage for tumor cell growth which ultimately leads to the transformation of normal cells into malignant tumor cells.[22] In the year 2000, Hanahan and Weinberg suggested six essential changes in cell physiology that together are the basis of cancer growth, namely self-sufficiency in growth signals, insensitivity to anti-growth signals, tissue invasion and metastasis, limitless replicative potential, sustained angiogenesis and evading apoptosis.[22] Later, it was proposed to also add immune hallmarks of cancer to this list.[23] In 2011 the concept of cancer hallmarks was reviewed and additional changes were added, namely genome instability and mutation, tumor promoting inflammation and also deregulating cellular energetics and avoiding immune destruction.[21] The involvement of glycosylation in these various hallmarks has been overviewed by Pearce et al., showing that different types of glycan groups play a role in cancer development.[24] It should be noted that the hallmarks of cancer are not exclusively true for cancer, but many of them can for example also apply to benign tumors which are unlikely to become malignant.[25]

In biomarker research the overlap of physiological processes in different diseases is often a challenge as it can highly influence the performance of a potential clinical test. In this thesis (Chapter 5) it was already discussed that inflammatory glycosylation changes were found in pancreatic cancer patients, which was explained by cancer related inflammation of the tumor microenvironment.[26] Here, it was also noted that the specificity of the analysis should be further evaluated in a cohort setting containing other cancer types and benign diseases, as this would better mimic a clinical setting. In other words, care has to be taken when drawing conclusions on specificity from a pure case-control study. In addition, as blood travels through the whole body and the sample is thus not locally derived at the tumor, it might be challenging to pinpoint where the alterations of blood-derived analytes are originating from.

In 2018, a prime example was presented that combined multiple types of blood-derived biomarkers (genetic- and protein markers) to identify the presence of cancer and the organ it originated from.[27] Unfortunately, this test appeared to be very sensitive for some cancers (ovary and liver), but for example for breast cancer the sensitivity of the test was poor.[27] Nevertheless, the concept of combining multiple analytes to increase sensitivity and specificity of blood-tests for cancer diagnosis was shown to be effective and this might be the way to go in the future. To this end, glycosylation is a valuable add-on to the genetic- and protein markers as was exemplified (Chapter 5) with profiles from pancreatic cancer patients that significantly differed from those of healthy controls.

Besides the use of these glycosylation changes as biomarkers, attention should also be paid to the origin of these changes. Understanding should be gained

on why the glycosylation of proteins in blood is changing and if this change serves a function in the progression of the disease. From cell-surface glycosylation it is known that it has a function in tumor development, but for secreted proteins this is not well understood.[28–30]

Concluding remarks

In summary, this thesis provides a high-throughput analysis method for released *N*-glycans including enzymatic release and automated sialic acid linkage-specific derivatization, cotton HILIC purification and MALDI-target spotting, for analysis of samples with high resolution MALDI-FTICR-MS (Chapter 2). It also describes the extension of this method for dried blood spots (Chapter 3) and approaches to further increase resolution to resolve overlapping species (Chapter 4). The automated method was applied to a pancreatic cancer study (Chapter 5), where glycosylation profiles were obtained from total serum and classification analysis was performed. A similar method was used for the total serum *N*-glycosylation analysis of breast cancer cases and controls (Chapter 6).

To translate biomarkers into the clinic, the analytical- and clinical validity of a test should always be critically evaluated. Samples should be collected in a standardized manner and (pre-)analytical protocols should be unambiguous in use. In addition, automation of methods and data analysis software can be very useful in terms of fast processing and repeatability, but critical steps of the procedures should always be critically monitored to ensure the quality of the conclusions drawn from the data. General glycosylation analysis from blood-derived samples can give a relatively simple and fast overview of the glycosylation changes with disease, but has its limitations in terms of origin of the glycans and isomer analysis. Therefore, the use of complementary analytical techniques is recommended for studies following total plasma/serum *N*-glycome analysis.

In the past, only limited number of biomarkers were successfully translated into the clinic, which might be the result of the complexity of the diseases, technical challenges and factors distracting the focus of the clinical needs. In future, much progress on overcoming these challenges should be made. In addition, overlapping physiological changes between diseases challenge the specificity of tests. Here, combining multiple types of analytes in one test might enable us in the future to better distinguish between different diseases and might ultimately lead to a blood test for the diagnosis of different types of cancer.

References

(1) Bladergroen, M. R.; Reiding, K. R.; Hipgrave Ederveen, A. L.; Vreeker, G. C. M.; Clerc, F.; Holst, S.; Bondt, A.; Wuhrer, M.; van der Burgt, Y. E. M. Automation of High-Throughput Mass Spectrometry-Based Plasma N -Glycome Analysis with Linkage-Specific Sialic Acid Esterification. J. Proteome Res. 2015, 14 (9), 4080–4086. https://doi.org/10.1021/acs.jproteome.5b00538.

(2) Selman, M. H. J.; Hemayatkar, M.; Deelder, A. M. A. M.; Wuhrer, M. Cotton HILIC SPE Microtips for Microscale Purification and Enrichment of Glycans and Glycopeptides. Anal. Chem. 2011, 83 (7), 2492–2499. https://doi.org/10.1021/ac1027116.

(3) Reiding, K. R.; Blank, D.; Kuijper, D. M.; Deelder, A. M.; Wuhrer, M. High-Throughput Profiling of Protein N-Glycosylation by MALDI-TOF-MS Employing Linkage-Specific Sialic Acid Esterification. Anal. Chem. 2014, 86 (12), 5784–5793. https://doi.org/10.1021/ac500335t.

(4) Reiding, K. R.; Lonardi, E.; Hipgrave Ederveen, A. L.; Wuhrer, M. Ethyl Esterification for MALDI-MS Analysis of Protein Glycosylation. Methods Mol. Biol. 2016, 1394, 151–162. https://doi.org/10.1007/978-1-4939-3341-9_11.

(5) Plummer, T. H.; Elder, J. H.; Alexander, S.; Phelan, A. W.; Tarentino, A. L. Demonstration of Peptide:N-Glycosidase F Activity in Endo-β-N-Acetylglucosaminidase F Preparations. J. Biol. Chem. 1984, 259 (17), 10700–10704.

(6) Vreeker, G. C. M.; Nicolardi, S.; Bladergroen, M. R.; van der Plas, C. J.; Mesker, W. E.; Tollenaar, R. A. E. M.; van der Burgt, Y. E. M.; Wuhrer, M. Automated Plasma Glycomics with Linkage-Specific Sialic Acid Esterification and Ultrahigh Resolution MS. Anal. Chem. 2018, 90 (20), 11955–11961. https://doi.org/10.1021/acs.analchem.8b02391.

(7) Vreeker, G. C. M.; Nicolardi, S.; Madunic, K.; Kotsias, M.; van der Burgt, Y. E. M.; Wuhrer, M. O- and N-Glycosylation Analysis of Cell Lines by Ultrahigh Resolution MALDI-FTICR-MS. Int. J. Mass Spectrom. 2020, 448, 116267. https://doi.org/10.1016/j.ijms.2019.116267.

(8) Lim, M. D.; Dickherber, A.; Compton, C. C. Before You Analyze a Human Specimen, Think Quality, Variability, and Bias. Anal. Chem. 2011, 83 (1), 8–13. https://doi.org/10.1021/ac1018974.

(9) De Noo, M. E.; Tollenaar, R. A. E. M.; Özalp, A.; Kuppen, P. J. K.; Bladergroen, M. R.; Eilers, P. H. C.; Deelder, A. M. Reliability of Human Serum Protein Profiles Generated with C8 Magnetic Beads Assisted Maldi-TOF Mass Spectrometry. Anal. Chem. 2005, 77 (22), 7232–7241. https://doi.org/10.1021/ac050571f.

(10) Dědová, T.; Grunow, D.; Kappert, K.; Flach, D.; Tauber, R.; Blanchard, V. The Effect of Blood Sampling and Preanalytical Processing on Human N-Glycome. PLoS One 2018, 13 (7), e0200507. https://doi.org/10.1371/journal.pone.0200507.

(11) Lee, D. H.; Kim, J. W.; Jeon, S. Y.; Park, B. K.; Han, B. G. Proteomic Analysis of the Effect of Storage Temperature on Human Serum. Ann. Clin. Lab. Sci. 2010, 40 (1), 61–70.

(12) Jansen, B. C.; Reiding, K. R.; Bondt, A.; Hipgrave Ederveen, A. L.; Palmblad, M.; Falck, D.; Wuhrer, M. MassyTools: A High-Throughput Targeted Data Processing Tool for Relative Quantitation and Quality Control Developed for Glycomic and Glycoproteomic MALDI-MS. J. Proteome Res. 2015, 14 (12), 5088–5098. https://doi.org/10.1021/acs.jproteome.5b00658.

(13) Patiny, L.; Borel, A. ChemCalc: A Building Block for Tomorrow's Chemical Infrastructure. J. Chem. Inf. Model. 2013, 53 (5), 1223–1228. https://doi.org/10.1021/ci300563h.

(14) Clerc, F.; Reiding, K. R.; Jansen, B. C.; Kammeijer, G. S. M.; Bondt, A.; Wuhrer, M. Human Plasma Protein N-Glycosylation. Glycoconj. J. 2016, 33 (3), 309–343. https://doi.org/10.1007/s10719-015-9626-2.

(15) Ioannidis, J. P. A. Biomarker Failures. Clin. Chem. 2013, 59 (1), 202–204. https://doi.org/10.1373/clinchem.2012.185801.

(16) Kirwan, A.; Utratna, M.; O'Dwyer, M. E.; Joshi, L.; Kilcoyne, M. Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. Biomed Res. Int. 2015, 2015, 490531. https://doi.org/10.1155/2015/490531.

(17) Diamandis, E. P. The Failure of Protein Cancer Biomarkers to Reach the Clinic: Why, and What Can Be Done to Address the Problem? BMC Med. 2012, 10 (1), 87. https://doi.org/10.1186/1741-7015-10-87.

(18) Kailemia, M. J.; Park, D.; Lebrilla, C. B. Glycans and Glycoproteins as Specific Biomarkers for Cancer. Anal. Bioanal. Chem. 2017, 409 (2), 395–410. https://doi.org/10.1007/s00216-016-9880-6.

(19) Burke, W. Genetic Tests: Clinical Validity and Clinical Utility. Curr. Protoc. Hum. Genet. 2014, 81 (1). https://doi.org/10.1002/0471142905.hg0915s81.

(20) Monaghan, P. J.; Lord, S. J.; St John, A.; Sandberg, S.; Cobbaert, C. M.; Lennartz, L.; Verhagen-Kamerbeek, W. D. J.; Ebert, C.; Bossuyt, P. M. M.; Horvath, A. R. Biomarker Development Targeting Unmet Clinical Needs. Clin. Chim. Acta 2016, 460, 211–219. https://doi.org/10.1016/j.cca.2016.06.037.

(21) Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: The next Generation. Cell 2011, 144 (5), 646–674. https://doi.

org/10.1016/j.cell.2011.02.013.

(22) Hanahan, D.; Weinberg, R. A. The Hallmarks of Cancer. Cell 2000, 100 (1), 57–70. https://doi.org/10.1016/S0092-8674(00)81683-9.

(23) Cavallo, F.; De Giovanni, C.; Nanni, P.; Forni, G.; Lollini, P. L. 2011: The Immune Hallmarks of Cancer. Cancer Immunol. Immunother. 2011, 60 (3), 319–326. https://doi.org/10.1007/s00262-010-0968-0.

(24) Pearce, O. M. T. Cancer Glycan Epitopes: Biosynthesis, Structure and Function. Glycobiology 2018, 28 (9), 670–696. https://doi.org/10.1093/glycob/cwy023.

(25) Lazebnik, Y. What Are the Hallmarks of Cancer? Nat. Rev. Cancer 2010, 10 (4), 232–233. https://doi.org/10.1038/nrc2827.

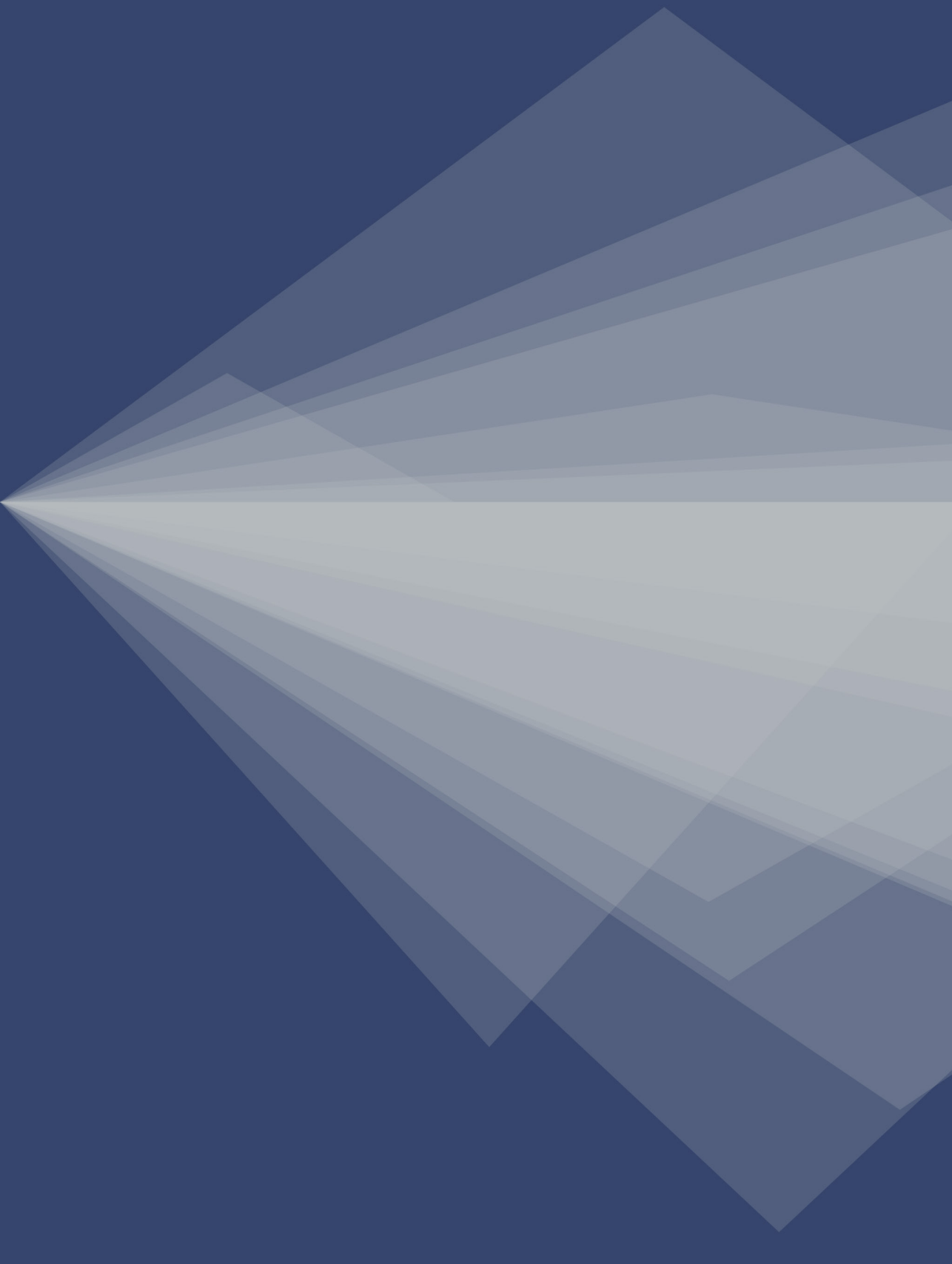
(26) Arnold, J. N.; Saldova, R.; Abd Hamid, U. M.; Rudd, P. M. Evaluation of the Serum N-Linked Glycome for the Diagnosis of Cancer and Chronic Inflammation. Proteomics 2008, 8 (16), 3284–3293. https://doi.org/10.1002/pmic.200800163.

(27) Cohen, J. D.; Li, L.; Wang, Y.; Thoburn, C.; Afsari, B.; Danilova, L.; Douville, C.; Javed, A. A.; Wong, F.; Mattox, A.; et al. Detection and Localization of Surgically Resectable Cancers with a Multi-Analyte Blood Test. Science (80-.). 2018, 359 (6378), 926–930. https://doi.org/10.1126/science.aar3247.

(28) Bassagañas, S.; Pérez-Garay, M.; Peracaula, R. Cell Surface Sialic Acid Modulates Extracellular Matrix Adhesion and Migration in Pancreatic Adenocarcinoma Cells. Pancreas 2014, 43 (1), 109–117. https://doi.org/10.1097/MPA.0b013e31829d9090.

(29) Schultz, M. J.; Swindall, A. F.; Bellis, S. L. Regulation of the Metastatic Cell Phenotype by Sialylated Glycans. Cancer Metastasis Rev. 2012, 31 (3–4), 501–518. https://doi.org/10.1007/s10555-012-9359-7.

(30) Lu, J.; Gu, J. Significance of β-Galactoside A2,6 Sialyltranferase 1 in Cancers. Molecules 2015, 20 (5), 7509–7527. https://doi.org/10.3390/molecules20057509.



Appendices

Abbreviations

ACN	acetonitrile
AGP	α-1-acid glycoprotein
Asn	asparagine
AUC	area under curve
BSM	bovine submaxillary glands
CA	carbohydrate antigen
CE	capillary electrophoresis
CEA	carcinoembryonic antigen
CEX	cation-exchange
CI	confidence intervals
CID	collision-induced dissociation
DBS	dried blood spot
DCC	*N,N′*-dicyclohexylcarbodiimide
DCIS	ductal carcinoma in situ
DHB	2,5-dihydrobenzoic acid
DIC	*N,N′*-diisopropylcarbodiimide
DNA	deoxyribonucleic acid
DPS	dried plasma spots
DTT	DL-Dithiothreitol
EDC	1-ethyl-3-(3-dimethylamino)propyl carbodiimide
ER	endoplasmic reticulum
ESI	electrospray ionization
FTICR	Fourier-transform ion cyclotron resonance
Fuc	fucose
FBS	fetal bovine serum
Gal	galactose
GalNAc	*N*-acetylgalactosamine
GlcNAc	*N*-acetylglucosamine
GuHCl	guanidine hydrochloride
HILIC	hydrophilic interaction liquid chromatography
HOBt	1-hydroxybenzotriazole
HT	high-throughput
IPMN	intraductal papillary mucinous neoplasms
IR	infrared

LC	liquid chromatography
LIF	laser-induced fluorescence detection
LLE	liquid-liquid extraction
LTR	lifetime risk
MALDI	matrix-assisted laser desorption/ionization
Man	mannose
MQ	milli-Q water
MRI	magnetic resonance imaging
MS	mass spectrometry
NET	neuroendocrine tumors
MTP	microtitration plate
Neu5Ac	*N*-acetylneuraminic acid
NMR	nuclear magnetic resonance
NP-40	nonidet P-40 substitute
OR	odds ratio
PDAC	pancreatic ductal adenocarcinomas
PGC	porous graphitized carbon
PNGase F	peptide-*N*-glycosidase F
PTM	post-translational modification
QC	quality control
RNA	ribonucleic acid
ROC	receiver operating characteristic
RP	resolving power *or* reversed phase
RSD	relative standard deviation
SD	standard deviation
SDS	sodium dodecyl sulfate
sLe	sialyl Lewis antigens
SPE	solid phase extraction
TFA	trifluoroacetic acid
TLC	thin layer chromatography
TOF	time of flight
TSNG	total serum *N*-glycome

English Summary

Glycosylation profiling with mass spectrometry

Clinical assays on body fluids are routinely performed for diagnostic and therapeutic purposes. Although an indispensable tool, it is widely acknowledged that such measurements of biomolecule concentrations leave room for improvement with regard to sensitivity and specificity. Moreover, for many diseases a detection modality based on a blood or urine specimen is not yet available and new biomarkers are urgently needed. In the case of cancer early detection often improves survival. Due to a globally increasing incidence of disease the screening for cancer turns more important. For pancreatic cancer, magnetic resonance imaging is currently considered the gold standard for high-risk patient screening, which is a relatively expensive technique. For breast cancer screening purposes, mammography is routinely applied although the diagnostic performance of this technique is rather limited, with for example a low sensitivity for dense breast tissues and low specificity.

Molecular markers are foreseen as a valuable addition to the imaging screening techniques currently used. Consequently, many studies have focused on method development for exploring biomarkers aiming for early detection. In a joint effort clinicians and researchers have been searching such markers since the early days of omics. Especially blood-based markers are of interest due to low invasiveness and therefore limited risk for the patient. Both for breast cancer and pancreatic cancer multiple protein candidate markers have been reported, however for various reasons these have not found their way into the clinic yet. An alternative strategy for finding new markers involves the exploration of post-translational modifications (PTMs) on proteins, or more generally profiling proteoforms. One of these PTMs is protein glycosylation, and is known that the expression of glycosyltransferases is altered in cancer, which has stimulated the biomarker field to search for candidate markers that relate to differences in protein glycosylation.

In the first part of this thesis the current clinical tests for the detection of pancreatic- and breast cancer are described to provide more insight into the need for new diagnostic methods (**Chapter 1**). Additionally, in this chapter the basics of human glycosylation and the biosynthetic pathway of glycans are outlined. Besides the biosynthesis, also the biological roles and the known relation with cancer are described. Moreover, the basic principles of glycosylation analysis and the analytical methods, for example

derivatization and mass spectrometry (MS), applied in this thesis are elaborated on.

The second part of this thesis is focused on the development of methods for glycosylation analysis (**Chapter 2, 3 and 4**). A previously established high-throughput serum/plasma protein glycosylation analysis method was further developed (**Chapter 2**). Side-reactions that can occur during the derivatization reaction were minimized, which resulted in a less complex mass spectrum. Additionally, the purification protocol was adjusted and glycan profiles were measured with matrix-assisted laser desorption/ionization (MALDI) Fourier transform ion cyclotron resonance (FTICR)-MS. This ultrahigh resolution mass analyzer is a highly sensitive and accurate platform for detection and assignment of glycans. The optimized method was extended to enable dried blood spot analysis in addition to serum- or plasma analysis (**Chapter 3**). From a clinical perspective, dried blood spots are easier to collect and convenient for shipment. The influence of storage time and temperature were evaluated. Sampling of dried blood spots is a promising method for specimen collection for released glycan analysis. In the last chapter of the method development section, the analysis of overlapping glycan signals with MALDI-FTICR-MS was explored (**Chapter 4**). The analysis of mono- and disaccharide glycan structures is generally difficult with MALDI-MS, as these signals are often located in the mass range where many matrix signals are present. These small glycans are also difficult to detect with other analytical platforms, such as porous graphitized carbon liquid chromatography-MS. Ultrahigh resolution MALDI-FTICR-MS combined with absorption mode readout resolves overlapping isobaric signals and successfully identifies such small glycans in complex cell line samples.

In the third part of this thesis, the developed serum glycan profiling method (**Chapter 2**) was applied to study the glycosylation profiles of pancreatic cancer and a similar method was used for breast cancer patients in order to evaluate glycan signatures as potential biomarkers for the detection of these cancers (**Chapter 5 and 6**). Released *N*-glycan profiles from two independent cohorts, both consisting of serum from pancreatic cancer patients and healthy controls, were measured and evaluated for the difference in glycosylation between the cases and controls (**Chapter 5**). Here, increases in branching and fucosylation were found and additionally a shift in sialic acid linkage was seen. The discriminating performance of the classification model built with these glycosylation features was good, which indicates the potential of glycosylation for detection of pancreatic cancer. In the breast cancer study, the released *N*-glycan profiles from breast cancer patients and healthy control were compared (**Chapter 6**). Interestingly, only three glycans were found to differ between cases and controls and also findings reported in literature were not replicated in this

study. However, it is noted that the findings reported in literature are very diverse and often contradicting each other. It is therefore hypothesized that the absence of a strong glycomic difference between cases and controls might be explained by the heterogeneity of the disease.

In the last part of this thesis a general discussion on biomarker research is provided (**Chapter 7**). Here, standardization of (pre-)analytical methods and the use of software for data-processing and interpretation is discussed. Additionally, the difficulties of translational research and cancer biomarker research are reflected on. Although the currently applied single biomarkers are an attractive starting point due to their straightforward cut-off values, the combination of multiple markers might allow to better distinguish between different diseases. In the future, this might be helpful in the development of blood tests for the early diagnosis of cancer and other diseases.

Nederlandse Samenvatting

Profilering van glycosylering met behulp van massaspectrometrie

Een *biomarker* is iets dat gemeten kan worden om de biologische toestand van een persoon te kunnen bepalen. Biomarkers worden routinematig gemeten met behulp van klinische testen voor diagnostische en therapeutische doeleinden. Hoewel deze testen een belangrijk onderdeel vormen van klinische zorgpaden, wordt algemeen erkend dat dergelijke bepalingen van bijvoorbeeld biomolecuulconcentraties kunnen verbeteren ten aanzien van de gevoeligheid en specificiteit. Bovendien is er voor veel ziekten geen detectiemethode op basis van een bloed- of urinemonsters beschikbaar en daarom is er dringend behoefte aan nieuwe biomarkers. Vanwege een wereldwijd toenemende incidentie van kanker wordt screening voor deze ziekte belangrijker. In het geval van kanker verbetert vroege detectie vaak de kans op langere overleving. Voor de opsporing van alvleesklierkanker is kernspintomografie (MRI) momenteel de gouden standaard voor de opsporing van de ziekte in hoog-risicopatiënten, echter dit is een relatief dure techniek. Voor opsporing van borstkanker wordt mammografie toegepast, hoewel deze techniek niet optimaal is voor bijvoorbeeld dicht borstweefsel en er nog veel fout positieve uitslagen gegeven worden.

Moleculaire biomarkers worden gezien als een waardevolle aanvulling op de momenteel gebruikte screeningstechnieken. Daarom richten veel studies zich op de ontwikkeling van methoden voor het verkennen van biomarkers met het oog op vroege detectie. In een gezamenlijke inspanning hebben clinici en onderzoekers nieuwe markers gezocht sinds de begindagen van *omics*. Vooral biomarkers die gemeten kunnen worden in bloed zijn interessant, omdat bloedprikken relatief laag invasief is en daardoor het risico voor de patiënt beperkt is. Voor zowel borstkanker als alvleesklierkanker zijn er meerdere kandidaat biomarkers, in de vorm van eiwitten, beschreven in de literatuur, maar om verschillende redenen worden deze nog niet in de kliniek toegepast. Een alternatieve strategie voor het vinden van nieuwe markers betreft het in kaart brengen van post-translationele modificaties (PTM's) op eiwitten; of meer in het algemeen de profilering van eiwitvormen. Het is bekend dat bij kanker de expressie van glycosyltransferasen verandert. Glycosyltransferasen zijn enzymen die suikergroepen aan elkaar verbinden, welke op hun beurt verbonden kunnen worden aan bijvoorbeeld eiwitten of lipiden. Zulke veranderingen in transferase hoeveelheden heeft onderzoekers gestimuleerd om eiwitglycosylering te bestuderen in de zoektocht naar biomarkers voor kanker.

In het eerste deel van dit proefschrift worden de huidige klinische diagnostische methoden voor de detectie van pancreas- en borstkanker beschreven waardoor het duidelijk wordt dat er behoefte is aan eenvoudige nieuwe methoden om deze ziekten vroeger op te sporen (**Hoofdstuk 1**). Bovendien worden in dit hoofdstuk de basisprincipes van humane glycosylering en de biosynthese van glycanen beschreven. Daarnaast worden de biologische functies van glycanen en de veranderingen bij kanker beschreven. Tenslotte worden de basisprincipes van glycaananalyse en de analytische methoden, zoals derivatisering en massaspectrometrie (MS), uitgelegd.

Het tweede deel van dit proefschrift is gericht op de ontwikkeling van methoden voor glycaananalyse (**Hoofdstuk 2, 3 en 4**). Een eerder ontwikkelde analysemethode voor serum/plasma eiwitglycosylering werd verder geoptimaliseerd (**Hoofdstuk 2**). Nevenreacties die kunnen optreden tijdens de derivatiseringsreactie werden geminimaliseerd, wat resulteerde in een minder complex massaspectrum. Daarnaast werd het zuiveringsprotocol aangepast en werden glycaanprofielen gemeten met behulp van een hoge-resolutie MALDI-FTICR-MS. Deze MS met ultrahoge resolutie is momenteel het meest gevoelige en nauwkeurige platform voor detectie en toekenning van glycanen. Deze geoptimaliseerde methode werd uitgebreid om naast serum- of plasma-analyse de analyse van opgedroogde bloeddruppels mogelijk te maken (**Hoofdstuk 3**). Vanuit een klinisch perspectief zijn gedroogde bloeddruppels gemakkelijker te verzamelen en verzenden. De invloed van opslagtijd en temperatuur werd geëvalueerd en hieruit werd geconcludeerd dat een gedroogde bloeddruppel een veelbelovende manier van monsterverzameling voor glycaananalyse is. In het laatste hoofdstuk van de sectie over methode-ontwikkeling is de analyse van overlappende glycaansignalen met MALDI-FTICR-MS onderzocht (**Hoofdstuk 4**). Mono- en disacharide glycaanstructuren zijn over het algemeen moeilijk te analyseren met MALDI-MS, omdat deze signalen zich vaak in het massagebied bevinden waar veel matrixsignalen aanwezig zijn. Bovendien zijn deze kleine glycanen ook moeilijk te detecteren met andere analytische platforms. MALDI-FTICR-MS met ultrahoge resolutie wordt daarom in de absorptiemodus gebruikt om de overlappende isobare signalen los van elkaar zichtbaar te maken en deze kleine glycanen met succes te identificeren in complexe monsters van cellijnen.

In het derde deel van dit proefschrift is de ontwikkelde glycaanprofileringsmethode (**Hoofdstuk 2**) toegepast om de glycaanprofielen van alvleesklierkanker te bestuderen. Een vergelijkbare methode werd gebruikt voor monsters van borstkankerpatiënten, waarbij de glycaanprofielen geëvalueerd werden als potentiële biomarkers voor de opsporing van deze ziekte (**Hoofdstuk 5 en 6**). Glycaanprofielen van twee onafhankelijke cohorten, die serum van alvleesklierkankerpatiënten

en gezonde individuen bevatten, werden gemeten en de verschillen tussen de patiënten en gezonde personen werden bestudeerd (**Hoofdstuk 5**). Er werd een toename in glycaanvertakking en -fucosylering gevonden en bovendien werd een verschuiving in siaalzuurverbinding waargenomen. Een classificatiemodel dat gemaakt werd op basis van deze glycaaninformatie, laat een goed onderscheidend vermogen zien tussen patiënten en gezonde individuen en laat daarmee de potentie van glycaanveranderingen als biomarkers zien. In de borstkankerstudie werden de glycaanprofielen van borstkankerpatiënten en gezonde vrouwen vergeleken (**Hoofdstuk 6**). Interessant is dat slechts drie glycanen verschillen tussen patiënten en gezonde vrouwen en dat ook bevindingen uit de literatuur niet konden worden gerepliceerd in deze studie. Er wordt echter opgemerkt dat de bevindingen in de literatuur erg divers zijn en ook niet altijd met elkaar overeenstemmen. Er wordt daarom verondersteld dat de afwezigheid van een sterk verschil in glycaanprofielen tussen patiënten en gezonde vrouwen kan worden verklaard door de heterogeniteit van de ziekte.

In het laatste deel van dit proefschrift wordt een algemene discussie over biomarkeronderzoek gegeven (**Hoofdstuk 7**). Hier wordt de standaardisatie van (pre-)analysemethoden en het gebruik van software voor gegevensverwerking en interpretatie besproken. Tenslotte worden de uitdagingen in translationeel onderzoek en biomarkeronderzoek voor kanker bediscussieerd. Ondanks dat de momenteel veel toegepaste enkelvoudige biomarkers een aantrekkelijk uitgangspunt zijn vanwege hun duidelijke grenswaarden, kan de combinatie van meerdere markers het mogelijk maken om beter onderscheid te kunnen maken tussen verschillende ziekten. In de toekomst zou dit kunnen leiden tot de ontwikkeling van een bloedtest voor de vroege opsporing van kanker en andere ziekten.

Curriculum Vitae

Geertruida (Gerda) Cornelia Maria Vreeker was born on July 20th, 1992 in Hoorn, the Netherlands. After she finished primary and secondary school she started her bachelor Chemistry at Vrije Universiteit Amsterdam, the Netherlands. After successfully obtaining her bachelor degree in 2013, she continued her studies with the master Chemistry at Vrije Universiteit Amsterdam/University of Amsterdam (combined program) with a specialization in Analytical Sciences. Next to her studies, Gerda has been involved in promotion activities of the studies, tutored at practical courses and was a member of the Program Committee Chemistry for two years.

During her masters she performed two internships. The first was in 2014 at the Medicinal Chemistry department at Vrije Universiteit Amsterdam and together with corresponding courses, this internship was extra-curricular and thus added to a complete analytical chemistry master. To gain more insights in molecules and chemical reactions she performed her internship on a subject related to drug design and synthesis. Besides the performance of synthesis experiments, she was also able to use her background in analytical chemistry to fully purify and characterize the synthesized molecules with liquid chromatography (LC), thin layer chromatography (TLC), infrared (IR) spectroscopy, UV detection, nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS). At the end of 2014 she started her second internship at the Center for Proteomics and Metabolomics (CPM) at Leiden University Medical Center (LUMC), Leiden, the Netherlands. Here she focused on method development for glycosylation analysis and did a cohort analysis on an automated platform under the supervision of dr. Karli R. Reiding and prof. dr. Manfred Wuhrer.

In February 2016 Gerda started her PhD in a bridging position between the CPM and the department of Surgery (HLK) under the supervision of prof. dr. Manfred Wuhrer (CPM), prof. dr. Rob A. E. M. Tollenaar (HLK), dr. Wilma E. Mesker (HLK) and dr. Yuri E. M. van der Burgt (CPM). Her research project focused on the MS method development for the analysis of protein glycosylation from blood-based samples and the application of these methods to cancer biomarker studies.

List of publications (in this thesis)

- **Vreeker, G.C.M.**, Vangangelt, K.M.H., Bladergroen, M.R., Nicolardi, S., Mesker, W.E., Wuhrer, M., Van der Burgt, Y.E.M., Tollenaar, R.A.E.M.: Serum *N*-glycan profiles differ for various breast cancer subtypes. Glycoconj. J. (2021). doi:10.1007/s10719-021-10001-3

- **Vreeker, G.C.M.***, Hanna-Sawires, R.G.*, Mohammed, Y., Bladergroen, M.R., Nicolardi, S., Dotz, V., Nouta, J., Bonsing, B.A., Mesker, W.E., Van der Burgt, Y.E.M., Wuhrer, M.**, Tollenaar, R.A.E.M.**: Serum *N*-Glycome Analysis Reveals Pancreatic Cancer Disease Signatures. Cancer med. (2020). doi:10.1002/cam4.3439

- **Vreeker, G.C.M.**, Nicolardi, S., Madunic, K., Kotsias, M., Van der Burgt, Y.E.M., Wuhrer, M.: *O*- and *N*-glycosylation analysis from of cell lines by ultrahigh resolution MALDI-FTICR-MS. Int. J. Mass Spectrom. 448, 116267 (2020). doi:10.1016/j.ijms.2019.116267

- **Vreeker, G.C.M.**, Bladergroen, M.R., Nicolardi, S., Mesker, W.E., Tollenaar, R.A.E.M., van der Burgt, Y.E.M., Wuhrer, M.: Dried blood spot *N*-glycome analysis by MALDI mass spectrometry. Talanta. 205, 120104 (2019). doi:10.1016/j.talanta.2019.06.104

- **Vreeker, G.C.M.**, Nicolardi, S., Bladergroen, M.R., van der Plas, C.J., Mesker, W.E., Tollenaar, R.A.E.M., van der Burgt, Y.E.M., Wuhrer, M.: Automated Plasma Glycomics with Linkage-Specific Sialic Acid Esterification and Ultrahigh Resolution MS. Anal. Chem. 90, 11955–11961 (2018). doi:10.1021/acs.analchem.8b02391

* *These authors share first authorship*

** *These authors share last authorship*

List of publications (not in this thesis)

- Singh, S.S.*, Naber, A.*, Dotz, V., Schoep, E., Memarian, E., Slieker, R.C., Elders, P.J.M., **Vreeker, G.C.M.**, Nicolardi, S., Wuhrer, M., Sijbrands, E.J.G., Lieverse, A.G., 't Hart, L.M., van Hoek, M. Metformin and statin use associate with plasma protein *N*-glycosylation in people with type 2 diabetes. BMJ Open Diabetes Research and Care. 8, e001230 (2020). doi:10.1136/bmjdrc-2020-001230

- Hauwert, N.J., Mocking, T.A.M., Da Costa Pereira, D., Kooistra, A.J., Wijnen, L.M., **Vreeker, G.C.M.**, Verweij, E.W.E., De Boer, A.H., Smit, M.J., De Graaf, C., Vischer, H.F., de Esch, I.J.P., Wijtmans, M., Leurs, R.: Synthesis and Characterization of a Bidirectional Photoswitchable Antagonist Toolbox for Real-Time GPCR Photopharmacology. J. Am. Chem. Soc. 140, 4232–4243 (2018). doi:10.1021/jacs.7b11422

- Reiding, K.R., **Vreeker, G.C.M.**, Bondt, A., Bladergroen, M.R., Hazes, J.M.W., van der Burgt, Y.E.M., Wuhrer, M., Dolhain, R.J.E.M.: Serum Protein *N*-Glycosylation Changes with Rheumatoid Arthritis Disease Activity during and after Pregnancy. Front. Med. 4, 1–11 (2018). doi:10.3389/fmed.2017.00241

- **Vreeker, G.C.M.**, Wuhrer, M.: Reversed-phase separation methods for glycan analysis. Anal. Bioanal. Chem. 409, 359–378 (2017). doi:10.1007/s00216-016-0073-0

- Bladergroen, M.R., Reiding, K.R., Hipgrave Ederveen, A.L., **Vreeker, G.C.M.**, Clerc, F., Holst, S., Bondt, A., Wuhrer, M., van der Burgt, Y.E.M.: Automation of High-Throughput Mass Spectrometry-Based Plasma N -Glycome Analysis with Linkage-Specific Sialic Acid Esterification. J. Proteome Res. 14, 4080–4086 (2015). doi:10.1021/acs.jproteome.5b00538

ORCID iD: https://orcid.org/0000-0002-3709-929X

PhD Portfolio

PhD training

Mandatory courses

PhD Introductory Meeting	**2016**
Basic Methods and Reasoning in Biostatistics	**2016**
BROK Course	**2017**

Generic/disciplinairy courses

Quantitative LC-MS	**2016**
Summer course Glycosciences - *incl. poster presentation*	**2016**
Basic and Translational Oncology	**2016**
Communication in Science	**2018**
Clinical Diagnostics: Translational Mass Spectrometry	**2018**
Journal club on the chapters of "Essentials of Glycobiology"	**2018**
Introduction to GLP Regulations and Bioanalytical Method Validation by LC-MS/MS	**2019**

Attended lectures, LUMC presentations, participation in meetings

GlycoCan kickoff meeting	**2016**
Meeting Genootschap Keukenhof – *incl. presentation*	**2016, 2017**
LAP Career Event	**2019**
BCF Career Event	**2019**

Congress attendance and poster or oral presentations

CASA master event VU-UvA	**2016**
JP-NL Glycobiology meeting - *incl. poster*	**2016**
27th Glycobiology meeting Nijmegen - *incl. oral and poster*	**2016**
Symposium Clinical Mass Spectrometry Imaging	**2017**
American Society for Mass Spectrometry (ASMS) conference - *incl. oral*	**2018**
Nederlandse Vereniging voor Massaspectrometrie (NVMS) symposium - *incl. oral*	**2018**
NVMS conference attendance	**2019**
American Society for Mass Spectrometry (ASMS) conference - *incl. poster*	**2019**
Chains 2019 - *incl. poster*	**2019**

Other activities

Lab tour Rotary Club Noordwijk	**2016**
Lab tour Feadship - *incl. presentation*	**2017**
Clinical internship Oncology	**2019**

Teaching activities

Lecturing, lab assistance, student supervision

FOS course CPM – lab assistance	**2016**
FOS course Surgery - lecture	**2016, 2017, 2018, 2019**
Student supervision: Corné van der Plas	**2017**
Student supervision: Jennifer Heikamp	**2017**
Science4U project of Stedelijk Gymnasium Leiden	**2017, 2018, 2019**
FOS course Parasitology – lab assistance	**2019**

Parameters of esteem

Personal grants

NVMS conference attendance grant	**2017**

Organizational Experience

Social activities

Sinterklaas celebration CPM	**2016, 2017, 2018, 2019**
Borrel committee CPM	**2016, 2017**
Labouting CPM	**2016**
Christmas dinner CPM	**2017**

Acknowledgements

Here, at the end of this thesis I would like to thank all colleagues, friends and family for the support they gave me during my PhD.

I would like to express my gratitude towards Manfred Wuhrer, my first promotor. Thank you for your supervision, your fast feedback to manuscripts and your availability for discussions. Additionally, I would like to thank Rob Tollenaar, my second promotor, and Wilma Mesker, my co-promotor. Thank you for your supervision and support on the clinically oriented part of this thesis. Without your effort on the collection of samples and ambition to develop a blood test, this research would not have been possible.

Also, I would like to thank Yuri van der Burgt for the support in writing the manuscripts and completing revisions. I want to thank Marco Bladergroen for his technical support with the automated pipetting platform and the creation of many of the beautiful figures found in this thesis. For the technical support with the MALDI-FTICR-MS I would like to thank Simone Nicolardi.

In addition, I want to thank my paranymphs Jan Nouta and Annelies Vlaar and all (former) colleagues and students from both the Center for Proteomics and Metabolomics and the Surgery department for the collaborations and support you provided. A special thanks to all co-authors of the chapters in this thesis for their contributions to my thesis. Additionally, I would like to thank all colleagues I shared an office with for the great times we had and the support you gave me in the ups and downs of my PhD trajectory.

I would also like to express my sincere appreciation for all patients and healthy volunteers participating in this research. Without their contributions this research would not have been possible.

Additionally, I want to thank my family and friends for their support. Ik wil jullie bedanken voor de interesse die jullie in mijn werk getoond hebben de afgelopen jaren en voor de moeite die jullie genomen hebben om te proberen te begrijpen waar ik aan werkte.

Mam, bedankt voor alles wat u gedaan en gelaten hebt om mij de kansen te geven die ik gekregen heb en voor uw liefde en steun.

Also, a special thanks to my cats, Sammy and Sari. You make that even on the rainiest day there is always some sunshine coming from you.

Last, but definitely not least, I would like to express my gratefulness to my husband and best friend, Dennis van Dalen. Lieve Dennis, ik kom woorden te kort om jou te kunnen vertellen hoe dankbaar ik ben voor jouw steun, jouw vertrouwen, je adviezen, je geduld en bovenal je liefde. Bedankt dat je in hoogte- en dieptepunten naast me staat en mij de mooiste kanten van het leven laat zien.

Thank you,

Gerda van Dalen–Vreeker