Glyco(proteo)mic workflows for cancer biomarker discovery

Moran, A.B.

Chapter 1

Introduction

Introduction

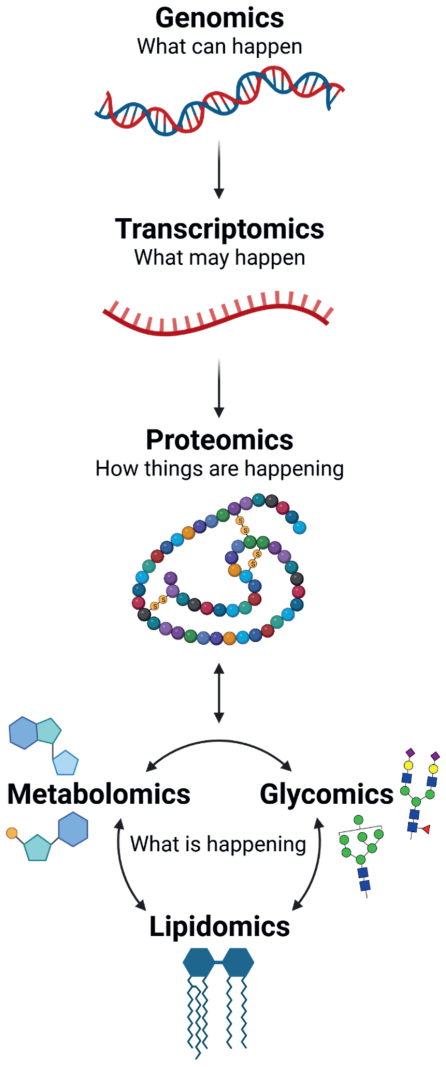
Cancer describes a vast set of diseases in which multiple alterations occur to normal healthy cells in order to transform them towards malignant derivatives.[1] These aberrations can have numerous and variable origins, occur in different types of cells as well as various genetic and metabolic pathways.[2] As a result, the presentation of the cancer phenotype can be diverse and, due to this complexity, the same clinical intervention may not work equally for different patients who seemingly suffer from the same condition.[3] As a result, a personalized approach to medicine may lead to improved outcomes for patients. In this sense, personalized medicine aims to use our understanding of the molecular pathways[4] manipulated by cancer in order to prescribe the right intervention at the right time to the right patient.[5]

Biomarker Discovery

Disease complexity may be resolved by following a set of biological signatures or "biomarkers" that are correlated with the disease and thus provide an individual perspective of the disease for each patient. A biomarker may be described as a biological parameter that is objectively measured to give information regarding biological, pathological or therapeutic processes.[6] To provide insights into these complex biological processes and interactions, the presence, absence, increase, or decrease of a specific biomarker, can be monitored. As a result, biomarkers may be used for clinical applications such as screening, diagnosis, prognosis, as well as the prediction of response to treatment and monitoring disease progression.[7] Thus, biomarkers have great importance and huge potential for improving patient health by means of developing non-invasive testing, greater clinical accuracy, and personalized medicine.

The discovery and development of biomarkers for clinical use is not limited to a single field and, as shown is **Scheme 1**, encompasses research into the expression of genes (genome), RNAs (transcriptome), proteins (proteome), and metabolites (metabolome)[8] as well as lipids (lipidome) and glycans (glycome). Although integrating multiple fields would provide a comprehensive overview of the biological system of interest,[8,9] the availability of resources, time and expertise pertains that researchers must often focus on a specific area. As physiological changes caused by disease are

often reflected in the proteome, this field represents an attractive opportunity for biomarker discovery and development.[10] In relation to this, proteins demonstrate great versatility and flexibility to carry out the functions of the cell, in contrast to the genome which is generally constant.



Scheme 1. Overview of the 'omics' fields. *Created with BioRender.com.*

Challenges for Biomarker Research

The chief requirement of a clinical biomarker is that it must "improve patient outcomes".[11] Moreover, it must be objectively measured using a suitable and analytically validated assay, which must also demonstrate sufficient performance in important areas such as clinical sensitivity and specificity, as well as positive and negative predictive value.[12,13] Thus, failure to meet these key performance metrics leads to challenges for biomarker testing whereby it is questionable whether patient outcomes are improved or not. A well-known example in this regard is the case of the prostate-specific antigen (PSA) test, an Food and Drug Administration (FDA)-approved test for early detection of prostate cancer (PCa).[14] The test is based primarily on the premise that the development of PCa is associated with the disruption of the basement membrane and basal cells of the prostate. As a result, an elevated PSA concentration is observed in serum after the protein "leaks" into the circulatory system.[15] However, it has been reported that the PSA test has a sensitivity of 21% and a specificity of 91% for detection of any PCa when a cut-off of 4.0 ng/mL is applied.[16] Although the sensitivity for high-grade cancer (Gleason score ≥ 8) was reported as 51% at this cut-off value.[16] As a result, most PCa will not be detected in patients (false-negative) using a PSA test alone. In addition, it has been shown that high-grade cancers are found amongst men with PSA levels < 4.0 ng/mL.[17] Thus, the test is unable to differentiate malignant and benign forms of the disease. Furthermore, an elevated PSA concentration in serum may also be the result of other factors and conditions, such as benign prostate hyperplasia (BPH).[18] Thus, when the cut-off is lowered to 3.0 ng/mL (32% sensitivity and 85% specificity for detection of any PCa), the likelihood of a false-positive diagnosis as well as overdiagnosis increases.[16] Overall, the poor performance of this biomarker in these important areas leads to underdiagnosis (false-negatives),[16] whereby the malignant disease is not detected early enough in order to trigger an effective clinical intervention, and overdiagnosis (false-positives).[19] As a result, patients with a benign or non-cancerous condition undergo an invasive and unnecessary biopsy procedure.

A single biomarker approach such as the PSA test is an appealing strategy as concentration cut-off values are straightforward to implement.[20] Evidently, however, there are pitfalls associated with this method which may arise due to a failure to take protein complexity into account. Proteome complexity mainly arises due to post-

transcriptional and post-translational events such as alternative splicing[21] as well as post-translational modifications (PTMs)[22] which may generate multiple proteoforms of the same protein that is transcribed from a single gene.[23,24] Thus, single biomarker assays, including many assays using antibody-based detection, generally do not distinguish between various proteoforms which may lead to inaccurate results.[25]

Opportunities for Discovery

The challenges posed by proteoforms for clinical testing also present future opportunities for biomarker discovery and development. It has been postulated that improvements in clinical test performance will be achieved when proteoforms are incorporated into clinical assays.[23,26] For example, pathophysiological changes associated with cancer cell development result in changes to various proteoforms, including PTMs such as glycosylation, in addition to differences in protein expression levels. Glycosylation is a critical PTM as correct glycosylation is required for protein folding, interactions between receptors and ligands, cellular signaling and recognition, and immune response.[27] Thus, oncogenesis is associated with aberrant glycosylation as the expression of regulatory enzymes such as glycosyltransferases as well as the availability of monosaccharide residues is impacted.[28–33] As a result, the interaction between the onset of cancer and glycosylation also represents potential opportunities for biomarker discovery and development.[30,31] In fact, the majority of FDA-approved cancer biomarkers currently implemented in clinical practice are glycoproteins,[34] although the glycosylation-specific features are often not taken into account during clinical testing.[35] In any case, this has given rise to the development of glycosylation-focused “omics” fields, namely the glycoproteome and glycome,[36] that may be explored in their own right as biomarkers of disease.

Glycosylation

Glycosylation refers to the process of modifying a carrier protein or lipid with carbohydrate molecules. In this process, sugar building blocks, also known as monosaccharides, are used to form larger and more complex carbohydrate structures, better known as oligosaccharides or glycans. These structures consist of a series of monosaccharides covalently linked together via glycosidic bonds. As shown in **Figure 1A**, there are seven monosaccharides that are commonly found throughout the human

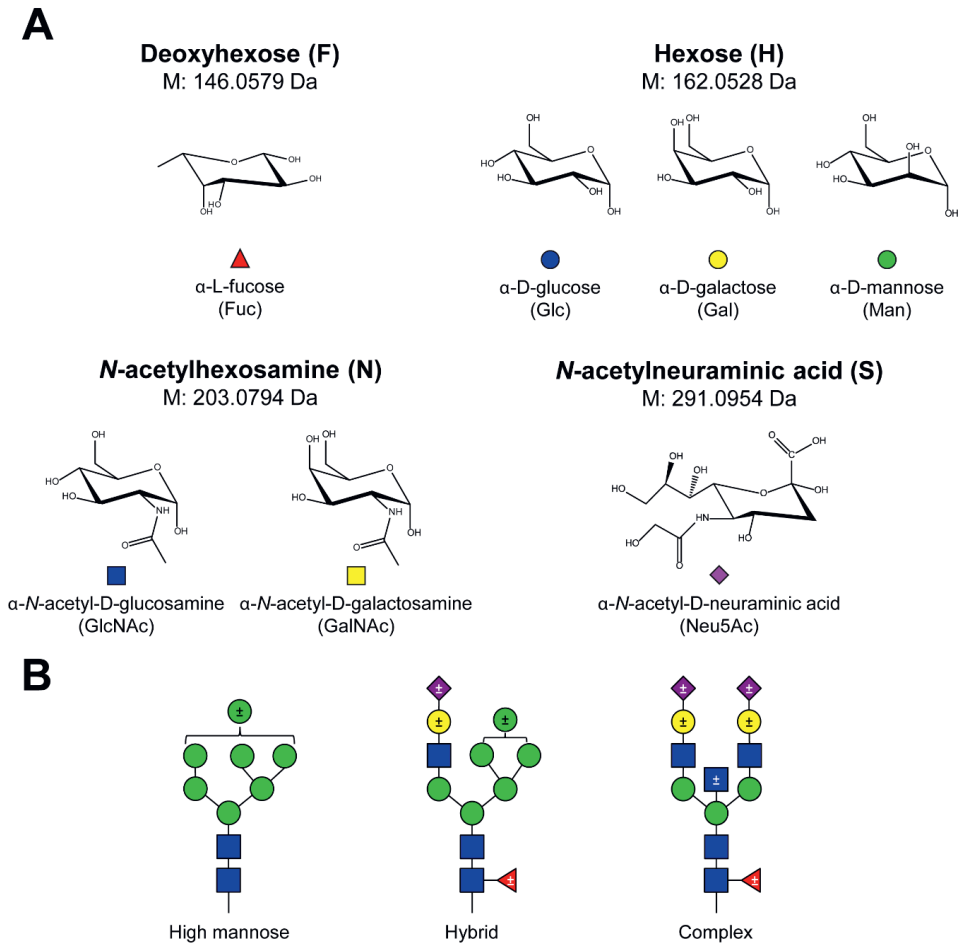


Figure 1. Most common monosaccharides found in humans. (A) Monosaccharides are shown in the alpha (α) conformation. Abbreviations for each group of sugars that share the same monoisotopic mass (M) as well as each individual sugar residue are shown. In this case, fucose (F) and sialic acid (S) are the primary deoxyhexose and *N*-acetylneuraminic acid found in humans, respectively. **(B)** The three glycans types, high mannose, hybrid and complex are shown. The symbol "±" denotes that monosaccharides may be further added or subtracted from the structure. The colored symbol for each monosaccharide is depicted according to the Consortium of Functional Glycomics.[102]

body as part of different glycoconjugates whereby the term "conjugate" refers to another compound, often a protein or a lipid, that is decorated with glycan structures. In the case of proteins, the glycans may be *N*- or *O*-linked. The former refers to the attachment of the glycan to the nitrogen atom of an asparagine (Asn) when the motif Asn-X-serine (Ser)/threonine (Thr) is found in the amino acid sequence whereby "X"

is any amino acid except proline (Pro). An *O*-linked glycan refers to glycans that are bound to the protein via the oxygen atom of Ser or Thr. The sequences of Asn-X-Ser/Thr and Ser/Thr are referred to as the glycosylation sites of the protein and there may be multiple *N*- and *O*-linked glycosylation sites present on a single protein. Furthermore, different glycosylation sites may also have different occupation levels. Thus, therein lies the complexity that is found on glycoproteins as macroheterogeneity (multiple glycosylation sites) and microheterogeneity (different glycans found on the same glycosylation site) contribute to the formation of different glycoforms of the same protein.[37]

In eukaryotes, glycosylation begins in the endoplasmic reticulum (ER). In order to produce *N*-linked glycoproteins, first a lipid-linked oligosaccharide is built by linking multiple mannoses and glucoses to a GlcNAc that is priorly attached to the lipid. Following this, the glycan portion is transferred to a polypeptide chain whereby it ensures proper folding during post-translation modification of the protein before the entire glycoconjugate is translocated to the Golgi apparatus. Depending on the location of the glycoprotein in the Golgi apparatus (cis-, medial-, or trans-) as well as the presence of different glycosyltransferases, exoglycosidases, and endoglycosidases, **Figure 1B** shows that different glycan structures may be produced (high mannose, hybrid, and complex) before the glycoprotein is secreted by the cell.[37]

Glycans are structurally more complex than other macromolecules such as proteins owing to their non-linear production and branching nature. For example, polysaccharides are built by linking a monosaccharide to any available hydroxyl oxygen of another sugar via its anomeric carbon. Furthermore, glycosidic linkages may be present in an alpha (α) or beta (β) configuration. As a result, this gives rise to structures that may contain different compositions (monosaccharides), linkages, configurations (α- or β-), and branching. Thus, the structural diversity found within glycans gives rise to their numerous functions and roles in biological systems, including molecular structure and organization, as well as intrinsic and extrinsic signaling and recognition.[37] For instance, glycans with the same composition may have different linkages or structures (isomers) that affect their function. Sialic acids may mask or allow access to an underlying galactose for binding depending on whether they are α2,6- or α2,3-linked.[38] For example, α2,6-linked sialylation allows interaction with the asialoglycoprotein receptor whereas the α2,3-linked variant blocks

this interaction.[39] In addition, a galactosylated α6-antenna, rather than the α3-antenna, increases monoclonal antibody FcγRIIIA binding affinity.[40] Finally, it has been shown that core-fucosylation plays a role in regulating the immune system[41–43] whereas antennary fucosylation may function more in mediating host–microbe interactions.[44,45]

Glycoproteins as Cancer Biomarkers

The importance of glycosylation is evident in the malignant and physiological changes that result in aberrant glycosylation profiles during cancer progression. It has been shown that not only abnormal glycosylation patterns are generated, but that glycosylation actively contributes to tumor proliferation, metastasis and angiogenesis.[46,47] As a result, aberrant glycosylation is continually being recognized as a hallmark of cancer. Undoubtedly there are a plethora of cancer-associated physiological alterations that affect cellular glycosyltransferases and glycosylation profiles, and vice versa. However, although glycosylation shows strong disease indications and greater promise for biomarker development, the glycosylation of human proteins and their clinical validity have not yet been thoroughly defined.[35] As a result, characteristic glycosylation signatures in cancer still require further investigation and, importantly, suitable methodology to measure these changes must be developed.

PSA Proteoforms as PCa Biomarkers

PSA is a glycoprotein with a single *N*-glycosylation (Asn_{69}) and, although the PSA test underperforms in relation to PCa screening and diagnosis, its various glycoforms have been associated with differentiating specific groups of PCa patients. In relation to this, specific PSA glycosylation signatures such as an increase in both α2,3-sialylation[48] and fucosylation[49] were correlated with the distinction between PCa and non-PCa groups, as well as normal and tumor-derived cells, respectively. Furthermore, other possibilities for biomarker development may be found by examining the disruption of the biological processing of PSA into its various proteoforms during prostate disease. **Figure 2** summarizes the PSA proteoforms that are found in serum during normal bioprocessing of the protein in healthy cells. The prepropeptide form of PSA is produced in epithelial cells where it is co-translationally cleaved at Ala_{17} into [-7]proPSA PSA,[50,51] an inactive precursor form that contains a 7-amino acid long

propeptide, proPSA ([-7]proPSA).[52] Cleavage of [-7]proPSA at Arg_{24} in the lumen results in the active form of the protein (PSA), which is normally secreted by the epithelial cells into semen. However, in cases of PCa, the epithelial lining of the prostate becomes disrupted, resulting in the entry of active PSA and inactive [-7]proPSA, as well as its constituent degradation products ([-5]proPSA, [-4]proPSA, [-

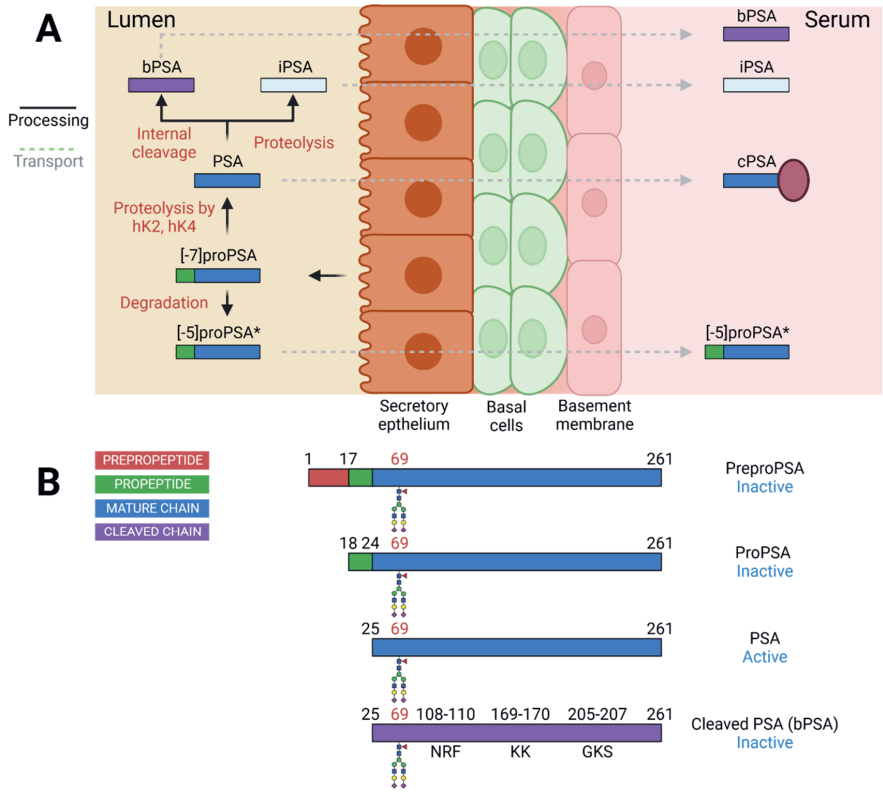


Figure 2. Normal bioprocessing of PSA. (A) The prepropeptide form of PSA is co-translationally cleaved into [-7]proPSA.[50,51] [-7]proPSA may also undergo further degradation which results in truncated proPSA forms, including [-5]proPSA. In the lumen.[50–52] [-7]proPSA undergoes proteolysis by hK2 and hK4 to obtain PSA. This active form of the protein is secreted by the prostate gland into seminal fluid, however it may also undergo further processing in the lumen to inactivate the protein, yielding internally cleaved PSA (bPSA) and inactive PSA (iPSA).[50,51] A small portion of each proteoform also enters into blood circulation whereby active PSA forms complexes with protease inhibitors which produces complexed PSA (cPSA).[50,51] Asterisk (*) refers to other proPSA forms such as [-4], [-3], [-2], and [-1]proPSA. **(B)** PSA proteoforms include preproPSA with signal peptide, [-7]proPSA, active PSA, and cleaved bPSA.[50] Amino acid positions are highlighted along the top. The *N*-glycosylation site is highlighted in red and the most abundant glycan (H5N4F1S2) is shown. The most prominent cleavage sites and associated amino acids are shown for bPSA.[50,51] *Created with BioRender.com.*

2]proPSA, [-1]proPSA) into the bloodstream.[50–52] Conversely, active PSA may also be inactivated by undergoing internal cleavage at specific amino acid positions (R_{109} – F_{110}, K_{169} – K_{170}, K_{206} –S_{207})[53–55] to produce inactive cleaved PSA[54], also known as benign PSA (bPSA)[56]. Interestingly, bPSA appears to be enriched in the transition zone of the prostate in hyperplastic BPH tissue which is thought to be a highly proteolytic environment.[56] Thus, the impact that disease has on the biological processing of PSA gives a clear example of how specific proteoforms may be investigated in order to improve the clinical performance of a diagnostic test when testing for PCa (proPSA) or BPH (bPSA). Despite these developments, comprehensive studies of PSA proteoforms, especially in body fluids other than serum, are lacking. Although PSA variants have been examined with some beneficial outcomes,[51] these tests are still dependent on protein concentration, rather than a proteoform profile. As a result, further studies into the clinical relevance of PSA proteoforms must still be carried out.

CRC *N*-Glycosylation Signatures

In relation to colorectal cancer (CRC), aberrant *N*-glycosylation signatures may be found during CRC progression. For example, a study that used CRC stem cells showed an increase in lectin binding for branched tri- and tetra-antennary structures, which was accompanied by an increase in the expression of *N*-acetylglucosaminyltransferase V (MGAT5),[57] the enzyme responsible for β1,6-branching.[58] Interestingly, MGAT5 overexpression is associated with an increase in malignancy and tumor growth rate in human colon carcinoma cells.[59] In addition, upregulation of ST6 beta-galactoside α-2,6-sialyltransferase 1 (ST6GAL1)[60,61] and a concurrent increase in expression of α2,6-sialylation is reported in CRC.[62] Here, sialylation modifies β1 integrin conformation, resulting in a protective effect against apoptosis in CRC cells.[63] Additionally, the aforementioned glycosylation signatures may also be investigated as biomarkers of the disease. For example, Vroome *et al.* showed an increase in tri- and tetra-antennary *N*-glycan structures as well as an overall elevation in α2,6-sialylation when comparing the total serum *N*-glycome of CRC patients with healthy controls.[64] Furthermore, the authors illustrated the discrimination of cases from controls and CRC survival was also predicted.[64] Another study focused more specifically on the glycosylation patterns of the current CRC biomarker, carcinoembryonic antigen (CEA).[65] CEA is a highly glycosylated protein

and it is expected that the assessment of its heterogeneous and abundant glycosylation could result in biomarkers that better reflect cancer progression.[66] In this case, Zhao *et al*. examined CEA from tumor tissues in comparison with paired tumor-adjacent normal tissues using a lectin microarray.[65] Greater levels of fucose and mannose were observed whilst there was a decrease in *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), galactose as well as branched and bisecting *N*-glycans on CEA.[65] Evidently, glycosylation is involved throughout numerous cancer processes in CRC, however further investigations are required in order to determine the most suitable approach (total serum *versus* specific glycoprotein), glycosylation signature, as well as analytical methodology that will enable translation of the results into the clinics.

The discovery and development of glycoproteins as biomarkers for clinical use requires suitable analytical methodology for this purpose. In this case, the technique must demonstrate sufficient sensitivity in order to detect low abundant analytes as well as specificity so that associations of specific analytes with the disease may be determined. In relation to the latter, antigen-binding approaches using antibody or lectin arrays are fraught with the potential for non-specific binding.[25,35] For example, Zhao *et al.* reported that *Aleuria aurantia* lectin (AAL) shows specificity for α1-2 (antennary) fucose[65] whereas other studies describe specificity towards α1-6 (core) fucose,[67] highlighting the ambiguity that may arise from such approaches. In contrast, measurands used in clinical tests must be clearly defined.[68] Furthermore, it is important to develop a proper understanding of the disease biology in order to develop biomarkers that can accurately monitor disease processes. Thus, it is important to build the required specificity into the biomarker test throughout the discovery and development period. In this sense, mass spectrometry (MS) is a technique that is ideally suited for this purpose as it allows specific identification via unique precursor mass and fragmentation patterns as well as direct quantification of proteins.[35,69]

Glyco(proteo)mics

Mass Spectrometry

MS describes a range of techniques wherein gas-phase ions of compounds are produced followed by the measurement of these ions as mass-to-charge ratios (*m/z*). The process of ionization produces gas-phase ions in the source, often via matrix-assisted laser desorption ionization (MALDI) or electrospray ionization (ESI). This is followed by mass separation of the ions in the mass analyzer under a vacuum. This is a critical function of MS as the ions are directly related to the nature, structure, and composition of the precursor molecule. Several types of mass analyzers have been developed for this purpose and, in many cases, they are combined in order to enhance the selectivity of the measurement. For example, quadrupole time-of-flight (QTOF) instruments implement a mass range pre-selection via the quadrupole which is followed by the separation of the ions inside a flight tube. The *m/z* of the ion is determined by the time taken to reach the detector, which is directly related to its mass. Finally, the data output from the detector is given as a mass spectrum, which may then be interpreted manually or using appropriate software in order to yield important information regarding the molecule of interest.[70]

The power of MS can be further enhanced by adding an extra separation dimension prior to mass detection. This may be achieved via the coupling of an online separation technique before the introduction of the sample into the mass spectrometer. With regard to the analysis of glycoproteins, two separation techniques have demonstrated a wide range of applications when hyphenated with MS: capillary electrophoresis (CE)–ESI–MS[66,71–77] and liquid chromatography (LC)–ESI–MS.[78–84]

Capillary Electrophoresis–Electrospray Ionization–Mass Spectrometry

Electrophoresis refers to the separation of compounds in solution following the application of voltage in order to generate an electric field.[85] Thus, analytes will migrate towards the cathode or anode depending on whether they are positively or negatively charged, respectively. In the case of capillary zone electrophoresis (CZE), the capillary is filled with a buffer known as the background electrolyte (BGE) in order to facilitate separation.[86] Depending on the charge of the analytes and their hydrodynamic size, each analyte displays its own electrophoretic mobility. In this case, an analytes charge

is determined based on its pKa as well as the pH, ionic strength, and concentration of the BGE.[87] However, the migration of each analyte within the capillary depends on the combined influence of its electrophoretic mobility as well as the electroosmotic flow (EOF) of the system. The EOF is generated by applying a voltage to the capillary and an interaction between cations, anions, and the capillary wall will occur, resulting in a bulk flow of the solution. In relation to this, the strength of the EOF may be influenced by controlling parameters such as the properties of the BGE, the coating present on the capillary internal wall, and the applied voltage.[87] Moreover, the resulting charge of the compound of interest and direction of the EOF must also be carefully considered to ensure that the analyte migrates efficiently towards the detector.

CE is regarded as an analytical technique which generates excellent peak resolution and high peak capacity, mainly because of the sharp peaks that are produced due to the flow profile that is EOF driven.[87] As a result, the hyphenation of CE with MS via ESI is a powerful approach which achieves high sensitivity and selectivity by the online separation of analytes based on their properties inside the capillary and used BGE. In the context of glycoproteomics, CE-ESI-MS has demonstrated efficient separation of sialylated glycans due to the negative charge that sialic acids carry.[88] Furthermore, isomeric separation can also be achieved as it was shown that sialic acids with different linkages are resolved as they have different pKas,[88] and likely also different hydrodynamic volumes, although the latter observation was not investigated. In addition, co-migrating compounds with different masses may be determined when coupled with MS detection. Overall, CE-ESI-MS has been applied across multiple areas in the glycomics and glycoproteomics fields, including the analysis of released glycans,[71–73] glycopeptides,[74,75] and intact glycoproteins.[76,77]

Liquid Chromatography–Electrospray Ionization–Mass Spectrometry

Chromatography achieves separation based on the interaction of compounds with the stationary and mobile phases. In this regard, the chemistry that is observed is largely dependent on the chromatographic system and the properties of the analyte inside that system. In relation to this, there are several chromatographic principles that are mainly applied in glycomic and glycoproteomics research, including high-pH/performance anion-exchange chromatography (HPAEC),[89] porous graphitized carbon (PGC)-LC,[78,83] reversed-phase (RP)-LC,[79,80] and hydrophilic interaction-LC

(HILIC).[81,90] In regard to HPAEC, glycans are separated following the application of a high pH which results in the deprotonation of monosaccharide hydroxyl groups.[91] In particular, this technique has been employed with pulsed-anomeric detection (PAD) for the analysis of individual monosaccharides following glycan hydrolysis.[89] Typically, this separation approach is difficult to couple with MS due to the incompatibility of the solvents, however hyphenation has been achieved via implementing an online desalting step prior to MS detection.[82] In general, the latter LC approaches (PGC, RP, HILIC) are more easily coupled with MS than HPAEC and have demonstrated numerous applications within the field of glycomics and glycoproteomics.[78–81,83,84,92] For example, hydrophobic and polar interactions influence PGC separation, as a result this technique demonstrates excellent separation of glycan isomers.[92,93] In the context of RPLC, a nonpolar stationary phase is employed and a noncovalent interaction is formed between the stationary phase and the analytes' nonpolar residues. Analytes elute in order of increasing hydrophobicity as their retention is dependent on the competitive solubilization between the stationary and mobile phases.[91] In the context of *N*-glycan analysis, elution is largely dominated by non-carbohydrate moieties at the reducing-end of the glycan.[91] Finally, HILIC is also a popular technique for separating glycoconjugates due to their hydrophilic nature and, as a result, separation is mainly determined by the glycans themselves. In this case, it has been reported that an analyte partitioning effect is created due to the presence of organic solvent in the mobile phase and the water-enriched layer at the polar stationary phase.[94,95] Thus, analyte elution is strongly influenced by the polar associations of the carbohydrate residues with the stationary phase, as well as other interactions which include hydrogen bonding, ion exchange and dipole-dipole interactions.[94] Therefore, HILIC-MS has been an important technique for the efficient separation and analysis of various types of glycoconjugates.

Top to Bottom MS Approaches

There are several approaches for analyzing glycoproteins by MS which range from top-down to bottom-up techniques (**Figure 3**). The approach for analyzing glycoproteins depends on the research question and level of detail that is required. For example, intact glycoprotein analysis involves the analysis of the entire glycoconjugate, with little sample preparation and, as a result, fewer modifications are introduced to the molecule during sample processing.[96] The separation may be

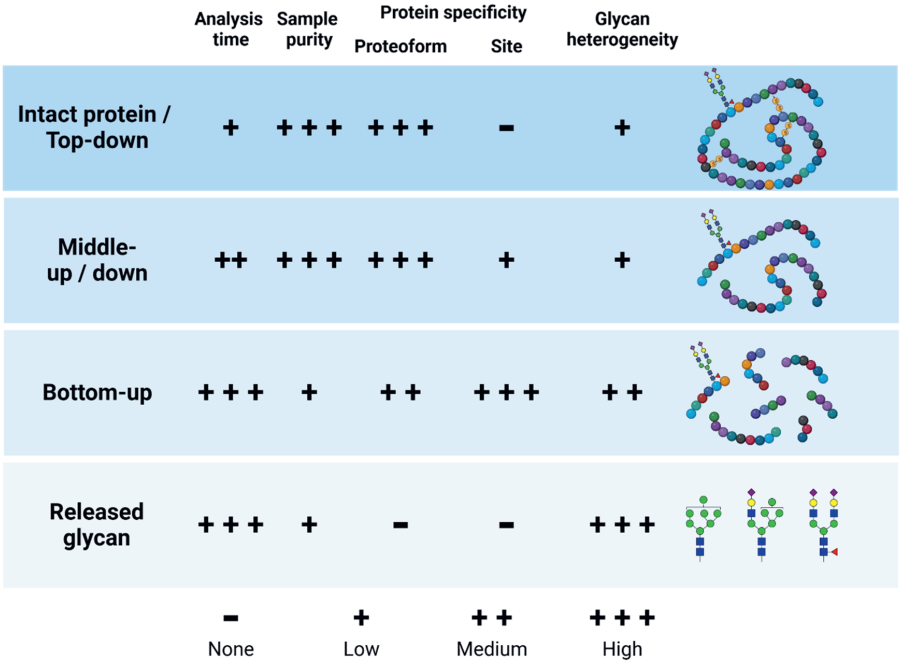


Figure 3. Top to bottom glyco(proteo)mic approaches. An ordinal scale is used to compare the four approaches across five categories, ranging from none to high. Analysis time refers to total time to perform sample preparation, measurement and data analysis. *Created with BioRender.com.*

performed under denaturing or native conditions with the aim of assessing the intact mass of different proteoforms as well as protein integrity and complexes when native conditions are employed.[97] Importantly, it should be noted that top-down analysis includes the performance of MS/MS fragmentation on the protein in order to gain sequence information.[97] However, intact protein or top-down analysis requires mainly isolated proteins and these techniques does not allow site-specific information to be gained. Furthermore, sensitivity is a challenge due to the dispersion of the signal over the charge envelope of the protein. Native top-down MS generates proteins with lower charge states,[98] however the aqueous solvent lacks the organic solvent and low pH that is often favorable for protein solubility and ion desolvation.[98]

Subunit analysis may also be carried out by following a middle-up or middle-down approach. Protein subunits may be generated by reducing protein disulfide bonds or enzymatic digestion, such as IdeS digestion of antibodies.[81] Similarly to the top-down approach mentioned above, the term “middle-down” is used when MS/MS

experiments are carried out during the measurement.[97] Several benefits are gained from following a protein subunit approach. For instance, the molecular weight of the analyte is lowered and, therefore, isotopic resolution may be achieved which allows the accurate mass to be determined. In addition, modifications may be assigned to specific subunits of the protein.[99] Despite these advantages, enzymes that cleave the protein into large polypeptides are generally not available for most proteins.[98]

Bottom-up analysis involves the enzymatic digestion of the protein and, unlike middle-up/down analysis, the protein is digested into peptides rather than subunits. Moreover, there is a large number of enzymes to choose from in order to generate smaller peptides, although trypsin is the most commonly applied.[100] This approach generates great detail about the glycoprotein of interest, including the identification of *N*-glycosylation sites and glycosylation site-occupancy, provided the peptide backbone is unique. Although glycopeptide analysis provides information regarding both the peptide and the glycan, the technique generates highly complex spectra. In addition, little information regarding the peptide sequence is gained from applying collisional-induced dissociation (CID), a commonly used technique for glycosidic bond fragmentation, and other fragmentation techniques should be utilized in this case.[100]

Finally, *N*-glycans may also be released entirely from the glycoprotein via enzymes such as an endoglycosidases. In this case, the complete range of glycan microheterogeneity may be assessed. This approach has several advantages, including the absence of any protein interference which produces spectra with less complexity. However, as a result, information regarding the carrier protein is lost. Furthermore, glycans possess poor ionization efficiency and do not contain a natural chromophore. In this case, sample preparation of glycans often involves reducing-end labeling with a fluorophore-containing compound in order to enable fluorescence detection.[101]

As demonstrated in **Figure 3**, each approach shows that there are several factors that should be considered when determining which tactic to follow, including the sample purity, time required for performing sample preparation and measurements, the modification of interest, and the complexity of the generated data. Alternatively, to gain a comprehensive picture, multiple approaches can be applied. For example, in cases where the assessment of different glycoprotein proteoforms via a top-down approach

may be the primary objective, bottom-up analysis would provide the necessary complimentary information regarding glycosylation site identification and occupancy determination. In addition, released glycan analysis can provide the full overview of glycan microheterogeneity and abundance. Moreover, subunit analysis provides further support regarding modifications to the protein sequence. Thus, a multilevel approach generates the most complete overview of the glycoprotein.

Scope

The aim of this thesis is to explore and develop MS workflows coupled with online separation for the discovery of glyco(proteo)mic signatures in cancer. In this case, two different approaches are taken in order to explore PCa and CRC, both in terms of sample processing and separation mode.

In the first part of this thesis, we focus on a specific glycoprotein, namely PSA, and perform intact protein analysis in order to assess the potential of its proteoforms for stratifying PCa patients. **Chapter 2** describes the development of a CE method in combination with MS detection in order to profile urinary intact PSA. In addition, the performance and application of the method is validated via intermediate and repeatability measurements as well as a small patient study. In **Chapter 3**, the data processing workflow is further developed, overcoming one of the main challenges that was identified in Chapter 2. Importantly, this chapter focuses on the same data set as Chapter 2 in order to validate the findings by comparing extracted ion and deconvolution quantification approaches. New insights are also obtained as the profiles of urinary and seminal PSA are compared, and further support for proteoform assignment is achieved via middle-up and bottom-up analyses.

In the second part of this thesis, the potential of total serum *N*-glycosylation as a biomarker for CRC is evaluated. **Chapter 4** describes the development and validation of a RPLC method in order to separate fluorescently labeled and sialic acid linkage derivatized *N*-glycans. Importantly, specific isomeric structures are differentiated by retention time and mass differences. Following this, the application of the method is demonstrated using a cohort of pre- and post-operative samples from CRC patient sera in **Chapter 5.** The clinical results are corroborated by comparison of the results with previous MALDI-MS measurements of the same cohort. Additional information is also gained by the separation of isomeric structures that was previously not observed.

Finally, a general discussion is provided in **Chapter 6**. Here, the analysis of isomeric *N*-glycans by RPLC is further examined. In addition, the technical challenges encountered during this thesis and potential improvements to the methodology are discussed. Finally, future perspectives regarding biomarker and method translation are described.

References

(1) Hanahan, D.; Weinberg, R. A. The Hallmarks of Cancer. *Cell*. **2000**, *100* (1). DOI: 10.1016/S0092-8674(00)81683-9.

(2) Fouad, Y. A.; Aanei, C. Revisiting the Hallmarks of Cancer. *American Journal of Cancer Research*. **2017**, *7* (5).

(3) Karapetis, C. S.; Khambata-Ford, S.; Jonker, D. J.; O'Callaghan, C. J.; Tu, D.; Tebbutt, N. C.; Simes, R. J.; Chalchal, H.; Shapiro, J. D.; Robitaille, S.; Price, T. J.; Shepherd, L.; Au, H.-J.; Langer, C.; Moore, M. J.; Zalcberg, J. R. K-Ras Mutations and Benefit from Cetuximab in Advanced Colorectal Cancer . *New England Journal of Medicine*. **2008**, *359* (17). DOI: 10.1056/nejmoa0804385.

(4) Chan, I. S.; Ginsburg, G. S. Personalized Medicine: Progress and Promise. *Annual Review of Genomics and Human Genetics*. **2011**, *12*. DOI: 10.1146/annurev-genom-082410-101446.

(5) Abrahams, E. Right Drug-Right Patient-Right Time: Personalized Medicine Coalition. *Clinical and Translational Science*. **2008**, *1* (1). DOI: 10.1111/j.1752-8062.2008.00003.x.

(6) Atkinson, A. J.; Colburn, W. A.; DeGruttola, V. G.; DeMets, D. L.; Downing, G. J.; Hoth, D. F.; Oates, J. A.; Peck, C. C.; Schooley, R. T.; Spilker, B. A.; Woodcock, J.; Zeger, S. L. Biomarkers and Surrogate Endpoints: Preferred Definitions and Conceptual Framework. *Clinical Pharmacology and Therapeutics*. **2001**, *69* (3). DOI: 10.1067/mcp.2001.113989.

(7) Borrebaeck, C. A. K. Precision Diagnostics: Moving towards Protein Biomarker Signatures of Clinical Utility in Cancer. *Nature Reviews Cancer*. **2017**, *17* (3). DOI: 10.1038/nrc.2016.153.

(8) Ilyin, S. E.; Belkowski, S. M.; Plata-Salamán, C. R. Biomarker Discovery and Validation: Technologies and Integrative Approaches. *Trends in Biotechnology*. **2004**, *22* (8). DOI: 10.1016/j.tibtech.2004.06.005.

(9) Nishizuka, S. et al. Diagnostic Markers That Distinguish Colon and Ovarian Adenocarcinomas: Identification by Genomic, Proteomic, and Tissue Array Profiling. *Cancer Research*. **2003**, *63* (17). DOI: 10.1097/00125480-200403000-00009.

(10) Zhan, X.; Long, Y.; Lu, M. Exploration of Variations in Proteome and Metabolome for Predictive Diagnostics and Personalized Treatment Algorithms: Innovative Approach and Examples for Potential Clinical Application. *Journal of Proteomics*. **2018**, *188*. DOI: 10.1016/j.jprot.2017.08.020.

(11) Holland, R. L. What Makes a Good Biomarker? *Advances in Precision Medicine*. **2016**, *1* (1). DOI: 10.18063/apm.2016.01.007.

(12) Shreffler, J.; Huecker, M. R. *Diagnostic Testing Accuracy: Sensitivity, Specificity, Predictive Values and Likelihood Ratios*; **2020**.

(13) Lalkhen, A. G.; McCluskey, A. Clinical Tests: Sensitivity and Specificity. *Continuing Education in Anaesthesia, Critical Care and Pain*. **2008**, *8* (6). DOI: 10.1093/bjaceaccp/mkn041.

(14) De Angelis, G.; Rittenhouse, H. G.; Mikolajczyk, S. D.; Blair Shamel, L.; Semjonow, A. Twenty Years of PSA: From Prostate Antigen to Tumor Marker. *Reviews in urology*. **2007**, *9* (3).

(15) Lilja, H.; Ulmert, D.; Vickers, A. J. Prostate-Specific Antigen and Prostate Cancer: Prediction, Detection and Monitoring. *Nature Reviews Cancer*. **2008**, *8* (4). DOI: 10.1038/nrc2351.

(16) Wolf, A. M. D.; Wender, R. C.; Etzioni, R. B.; Thompson, I. M.; D'Amico, A. V.; Volk, R. J.; Brooks, D. D.; Dash, C.; Guessous, I.; Andrews, K.; DeSantis, C.; Smith, R. A. American Cancer Society Guideline for the Early Detection of Prostate Cancer: Update 2010. *CA: A Cancer Journal for Clinicians*. **2010**, *60* (2). DOI: 10.3322/caac.20066.

(17) Thompson, I. M.; Pauler, D. K.; Goodman, P. J.; Tangen, C. M.; Lucia, M. S.; Parnes, H. L.; Minasian, L. M.; Ford, L. G.; Lippman, S. M.; Crawford, E. D.; Crowley, J. J.; Coltman, C. A. Prevalence of Prostate Cancer among Men with a Prostate-Specific Antigen Level ≤4.0 Ng per Milliliter. *New England Journal of Medicine*. **2004**, *350* (22). DOI: 10.1056/nejmoa031918.

(18) Monda, J. M.; Barry, M. J.; Oesterling, J. E. Prostate Specific Antigen Cannot Distinguish Stage T1a (A1) Prostate Cancer from Benign Prostatic Hyperplasia. *Journal of Urology*. **1994**, *151* (5). DOI: 10.1016/S0022-5347(17)35234-5.

(19) Carter, H. B. Prostate Cancers in Men with Low PSA Levels — Must We Find Them? *New England Journal of Medicine*. **2004**, *350* (22). DOI: 10.1056/nejme048003.

(20) van der Burgt, Y. E. M. Protein Biomarker Discovery Is Still Relevant and Has Entered a New Phase. *EBioMedicine*. **2019**, *43*. DOI: 10.1016/j.ebiom.2019.04.026.

(21) Blencowe, B. J. Alternative Splicing: New Insights from Global Analyses. *Cell*. **2006**, *126* (1). DOI: 10.1016/j.cell.2006.06.023.

(22) Prabakaran, S.; Lippens, G.; Steen, H.; Gunawardena, J. Post-Translational Modification: Nature's Escape from Genetic Imprisonment and the Basis for Dynamic Information Encoding. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*. **2012**, *4* (6). DOI: 10.1002/wsbm.1185.

(23) 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. *Clinics in Laboratory Medicine*. **2018**, *38* (3). DOI: 10.1016/j.cll.2018.05.001.

(24) Kelleher, N. L. A Cell-Based Approach to the Human Proteome Project. *Journal of the American Society for Mass Spectrometry*. **2012**, *23* (10). DOI: 10.1007/s13361-012-0469-9.

(25) Hoofnagle, A. N.; Wener, M. H. The Fundamental Flaws of Immunoassays and Potential Solutions Using Tandem Mass Spectrometry. *Journal of Immunological Methods*. **2009**, *347* (1–2). DOI: 10.1016/j.jim.2009.06.003.

(26) Kelleher, N. L.; Thomas, P. M.; Ntai, I.; Compton, P. D.; Leduc, R. D. Deep and Quantitative Top-down Proteomics in Clinical and Translational Research. *Expert Review of Proteomics*. **2014**, *11* (6). DOI: 10.1586/14789450.2014.976559.

(27) Lauc, G.; Rudan, I.; Campbell, H.; Rudd, P. M. Complex Genetic Regulation of Protein Glycosylation. *Molecular BioSystems*. **2010**, *6* (2). DOI: 10.1039/b910377e.

(28) Hakomori, S. Glycosylation Defining Cancer Malignancy: New Wine in an Old Bottle. *Proceedings of the National Academy of Sciences of the United States of America*. **2002**, *99* (16). DOI: 10.1073/pnas.172380699.

(29) Silva, M. L. S. Cancer Serum Biomarkers Based on Aberrant Post-Translational Modifications of Glycoproteins: Clinical Value and Discovery Strategies. *Biochimica et Biophysica Acta - Reviews on Cancer*. **2015**, *1856* (2). DOI: 10.1016/j.bbcan.2015.07.002.

(30) Reis, C. A.; Osorio, H.; Silva, L.; Gomes, C.; David, L. Alterations in Glycosylation as Biomarkers for Cancer Detection. *Journal of Clinical Pathology*. **2010**, *63* (4). DOI: 10.1136/jcp.2009.071035.

(31) Padler-Karavani, V. Aiming at the Sweet Side of Cancer: Aberrant Glycosylation as Possible Target for Personalized-Medicine. *Cancer Letters*. **2014**, *352* (1). DOI: 10.1016/j.canlet.2013.10.005.

(32) Vajaria, B. N.; Patel, P. S. Glycosylation: A Hallmark of Cancer? *Glycoconjugate Journal*. **2017**, *34* (2). DOI: 10.1007/s10719-016-9755-2.

(33) Munkley, J.; Elliott, D. J. Hallmarks of Glycosylation in Cancer. *Oncotarget*. **2016**, *7* (23). DOI: 10.18632/oncotarget.8155.

(34) Kirwan, A.; Utratna, M.; O'Dwyer, M. E.; Joshi, L.; Kilcoyne, M. Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. *BioMed Research International*. **2015**. DOI: 10.1155/2015/490531.

(35) de Haan, N.; Wuhrer, M.; Ruhaak, L. R. Mass Spectrometry in Clinical Glycomics: The Path from Biomarker Identification to Clinical Implementation. *Clinical Mass Spectrometry*. **2020**, *18*. DOI: 10.1016/j.clinms.2020.08.001.

(36) Hizal, D. B.; Wolozny, D.; Colao, J.; Jacobson, E.; Tian, Y.; Krag, S. S.; Betenbaugh, M. J.; Zhang, H. Glycoproteomic and Glycomic Databases. *Clinical Proteomics*. **2014**, *11* (1). DOI: 10.1186/1559-0275-11-15.

(37) Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H. *Essentials of Glycobiology, Third Edition*; **2017**.

(38) Schauer, R. Sialic Acids: Fascinating Sugars in Higher Animals and Man. *Zoology (Jena)*. **2004**, *107* (1), 49–64. DOI: 10.1016/j.zool.2003.10.002.

(39) Park, E. I.; Mi, Y.; Unverzagt, C.; Gabius, H. J.; Baenziger, J. U. The Asialoglycoprotein Receptor Clears Glycoconjugates Terminating with Sialic Acidα2,6GalNAc. *Proceedings of the National Academy of Sciences of the United States of America*. **2005**, *102* (47). DOI: 10.1073/pnas.0508537102.

(40) Hatfield, G.; Tepliakova, L.; Gingras, G.; Stalker, A.; Li, X.; Aubin, Y.; Tam, R. Y. Specific Location of Galactosylation in an Afucosylated Antiviral Monoclonal Antibody Affects Its FcγRIIIA Binding Affinity. *Frontiers in immunology*. **2022**, *13*, 972168. DOI: 10.3389/fimmu.2022.972168.

(41) Huhn, C.; Selman, M. H. J.; Ruhaak, L. R.; Deelder, A. M.; Wuhrer, M. IgG Glycosylation Analysis. *Proteomics*. **2009**, *9* (4), 882–913. DOI: https://doi.org/10.1002/pmic.200800715.

(42) Pučić, M. et al. High Throughput Isolation and Glycosylation Analysis of IgG-Variability and Heritability of the IgG Glycome in Three Isolated Human Populations. *Molecular and Cellular Proteomics*. **2011**, *10* (10). DOI: 10.1074/mcp.M111.010090.

(43) Hwang, H.; Jeong, H. K.; Lee, H. K.; Park, G. W.; Lee, J. Y.; Lee, S. Y.; Kang, Y. M.; An, H. J.; Kang, J. G.; Ko, J. H.; Kim, J. Y.; Yoo, J. S. Machine Learning Classifies Core and Outer Fucosylation of N-Glycoproteins Using Mass Spectrometry. *Scientific Reports*. **2020**, *10* (1). DOI: 10.1038/s41598-019-57274-1.

(44) Schneider, M.; Al-Shareffi, E.; Haltiwanger, R. S. Biological Functions of Fucose in Mammals. *Glycobiology*. **2017**, *27* (7). DOI: 10.1093/glycob/cwx034.

(45) Becker, D. J.; Lowe, J. B. Fucose: Biosynthesis and Biological Function in Mammals. *Glycobiology*. **2003**, *13* (7). DOI: 10.1093/glycob/cwg054.

(46) Fuster, M. M.; Esko, J. D. The Sweet and Sour of Cancer: Glycans as Novel Therapeutic Targets. *Nature Reviews Cancer*. **2005**, *5* (7). DOI: 10.1038/nrc1649.

(47) Pinho, S. S.; Reis, C. A. Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nature Reviews Cancer*. **2015**, *15* (9). DOI: 10.1038/nrc3982.

(48) Yoneyama, T.; Ohyama, C.; Hatakeyama, S.; Narita, S.; Habuchi, T.; Koie, T.; Mori, K.; Hidari, K. I. P. J.; Yamaguchi, M.; Suzuki, T.; Tobisawa, Y. Measurement of Aberrant Glycosylation of Prostate Specific Antigen Can Improve Specificity in Early Detection of Prostate Cancer. *Biochemical and Biophysical Research Communications*. **2014**. DOI: 10.1016/j.bbrc.2014.04.107.

(49) Peracaula, R.; Tabarés, G.; Royle, L.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M.; de Llorens, R. Altered Glycosylation Pattern Allows the Distinction between Prostate-Specific Antigen (PSA) from Normal and Tumor Origins. *Glycobiology*. **2003**, *13* (6). DOI: 10.1093/glycob/cwg041.

(50) Balk, S. P.; Ko, Y. J.; Bubley, G. J. Biology of Prostate-Specific Antigen. *Journal of Clinical Oncology*. **2003**, *21* (2). DOI: 10.1200/JCO.2003.02.083.

(51) Gilgunn, S.; Conroy, P. J.; Saldova, R.; Rudd, P. M.; O'Kennedy, R. J. Aberrant PSA Glycosylation - A Sweet Predictor of Prostate Cancer. *Nature Reviews Urology*. **2013**, *10* (2). DOI: 10.1038/nrurol.2012.258.

(52) Mikolajczyk, S. D.; Grauer, L. S.; Millar, L. S.; Hill, T. M.; Kumar, A.; Rittenhouse, H. G.; Wolfert, R. L.; Saedi, M. S. A Precursor Form of PSA (PPSA) Is a Component of the Free PSA in Prostate Cancer Serum. *Urology*. **1997**, *50* (5). DOI: 10.1016/S0090-4295(97)00449-4.

(53) Zhang, W. M.; Leinonen, J.; Kalkkinen, N.; Dowell, B.; Stenman, U. H. Purification and Characterization of Different Molecular Forms of Prostate-Specific Antigen in Human Seminal Fluid. *Clinical Chemistry*. **1995**, *41* (11). DOI: 10.1093/clinchem/41.11.1567.

(54) Mattsson, J. M.; Valmu, L.; Laakkonen, P.; Stenman, U. H.; Koistinen, H. Structural Characterization and Anti-Angiogenic Properties of Prostate-Specific Antigen Isoforms in Seminal Fluid. *Prostate*. **2008**, *68* (9), 945–954. DOI: 10.1002/pros.20751.

(55) Watt, K. W. K.; Lee, P. J.; M'Timkulu, T.; Chan, W. P.; Loor, R. Human Prostate-Specific Antigen: Structural and Functional Similarity with Serine Proteases. *Proceedings of the*

National Academy of Sciences of the United States of America. **1986**, *83* (10). DOI: 10.1073/pnas.83.10.3166.

(56) Mikolajczyk, S. D.; Millar, L. S.; Wang, T. J.; Rittenhouse, H. G.; Wolfert, R. L.; Marks, L. S.; Song, W.; Wheeler, T. M.; Slawin, K. M. "BPSA," a Specific Molecular Form of Free Prostate-Specific Antigen, Is Found Predominantly in the Transition Zone of Patients with Nodular Benign Prostatic Hyperplasia. *Urology*. **2000**, *55* (1). DOI: 10.1016/S0090-4295(99)00372-6.

(57) Guo, H.; Nagy, T.; Pierce, M. Post-Translational Glycoprotein Modifications Regulate Colon Cancer Stem Cells and Colon Adenoma Progression in Apcmin/+ Mice through Altered Wnt Receptor Signaling. *Journal of Biological Chemistry*. **2014**, *289* (45). DOI: 10.1074/jbc.M114.602680.

(58) 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). DOI: 10.18632/oncotarget.6283.

(59) Kim, Y. S.; Ahn, Y. H.; Song, K. J.; Kang, J. G.; Lee, J. H.; Jeon, S. K.; Kim, H. C.; Yoo, J. S.; Ko, J. H. Overexpression and β-1,6-N-Acetylglucosaminylation-Initiated Aberrant Glycosylation of TIMP-1: A "Double Whammy" Strategy in Colon Cancer Progression. *Journal of Biological Chemistry*. **2012**, *287* (39). DOI: 10.1074/jbc.M112.370064.

(60) Dall'Olio, F.; Chiricolo, M.; Mariani, E.; Facchini, A. Biosynthesis of the Cancer-Related Sialyl-Alpha 2,6-Lactosaminyl Epitope in Colon Cancer Cell Lines Expressing Beta-Galactoside Alpha 2,6-Sialyltransferase under a Constitutive Promoter. *European journal of biochemistry*. **2001**, *268* (22), 5876–5884. DOI: 10.1046/j.0014-2956.2001.02536.x.

(61) Dall'Olio, F.; Malagolini, N.; Chiricolo, M. Beta-Galactoside Alpha2,6-Sialyltransferase and the Sialyl Alpha2,6-Galactosyl-Linkage in Tissues and Cell Lines. *Methods in molecular biology (Clifton, N.J.)*. **2006**, *347*.

(62) Park, J. J.; Lee, M. Increasing the α 2, 6 Sialylation of Glycoproteins May Contribute to Metastatic Spread and Therapeutic Resistance in Colorectal Cancer. *Gut and Liver*. **2013**, *7* (6). DOI: 10.5009/gnl.2013.7.6.629.

(63) Zhuo, Y.; Chammas, R.; Bellis, S. L. Sialylation of Beta1 Integrins Blocks Cell Adhesion to Galectin-3 and Protects Cells against Galectin-3-Induced Apoptosis. *The Journal of biological chemistry*. **2008**, *283* (32), 22177–22185. DOI: 10.1074/jbc.M8000015200.

(64) 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. DOI: 10.18632/oncotarget.25753.

(65) Zhao, Q.; Zhan, T.; Deng, Z.; Li, Q.; Liu, Y.; Yang, S.; Ji, D.; Li, Y. Glycan Analysis of Colorectal Cancer Samples Reveals Stage-Dependent Changes in CEA Glycosylation Patterns. *Clinical Proteomics*. **2018**, *15* (1). DOI: 10.1186/s12014-018-9182-4.

(66) Pont, L.; Kuzyk, V.; Benavente, F.; Sanz-Nebot, V.; Mayboroda, O. A.; Wuhrer, M.; Lageveen-Kammeijer, G. S. M. Site-Specific N-Linked Glycosylation Analysis of Human Carcinoembryonic Antigen by Sheathless Capillary Electrophoresis-Tandem Mass Spectrometry. *Journal of Proteome Research*. **2021**, *20* (3). DOI: 10.1021/acs.jproteome.0c00875.

(67) Kekki, H.; Peltola, M.; van Vliet, S.; Bangma, C.; van Kooyk, Y.; Pettersson, K. Improved Cancer Specificity in PSA Assay Using Aleuria Aurantia Lectin Coated Eu-Nanoparticles for Detection. *Clinical Biochemistry*. **2017**, *50* (1–2). DOI: 10.1016/j.clinbiochem.2016.06.015.

(68) Smit, N. P. M.; Ruhaak, L. R.; Romijn, F. P. H. T. M.; Pieterse, M. M.; Van Der Burgt, Y. E. M.; Cobbaert, C. M. The Time Has Come for Quantitative Protein Mass Spectrometry Tests That Target Unmet Clinical Needs. *Journal of the American Society for Mass Spectrometry*. **2021**, *32* (3). DOI: 10.1021/jasms.0c00379.

(69) Matthiesen, R.; Bunkenborg, J. Introduction to Mass Spectrometry-Based Proteomics. In *Methods in Molecular Biology*; **2013**; Vol. 1007. DOI: 10.1007/978-1-62703-392-3_1.

(70) de Hoffmann, E.; Stroobant, V. *Mass Spectrometry: Principles and Applications*, 3rd ed.; Wiley: Chichester, **2007**.

(71) Marie, A.-L.; Ray, S.; Ivanov, A. R. Highly-Sensitive Label-Free Deep Profiling of N-Glycans Released from Biomedically-Relevant Samples. *Nature Communications*. **2023**, *14* (1), 1618. DOI: 10.1038/s41467-023-37365-4.

(72) Marie, A. L.; Ray, S.; Lu, S.; Jones, J.; Ghiran, I.; Ivanov, A. R. High-Sensitivity Glycan Profiling of Blood-Derived Immunoglobulin G, Plasma, and Extracellular Vesicle Isolates with Capillary Zone Electrophoresis-Mass Spectrometry. *Analytical Chemistry*. **2021**, *93* (4). DOI: 10.1021/acs.analchem.0c03102.

(73) 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. *Nature Communications*. **2019**, *10* (1), 1–8. DOI: 10.1038/s41467-019-09910-7.

(74) Giménez, E.; Ramos-Hernan, R.; Benavente, F.; Barbosa, J.; Sanz-Nebot, V. Analysis of Recombinant Human Erythropoietin Glycopeptides by Capillary Electrophoresis Electrospray-Time of Flight-Mass Spectrometry. *Analytica Chimica Acta*. **2012**, *709*. DOI: 10.1016/j.aca.2011.10.028.

(75) 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. *Analytical Chemistry*. **2018**, *90* (7), 4414–4421. DOI: 10.1021/acs.analchem.7b04281.

(76) Gstöttner, C.; Nicolardi, S.; Haberger, M.; Reusch, D.; Wuhrer, M.; Domínguez-Vega, E. Intact and Subunit-Specific Analysis of Bispecific Antibodies by Sheathless CE-MS. *Analytica Chimica Acta*. **2020**, *1134*. DOI: 10.1016/j.aca.2020.07.069.

(77) Moran, A. B.; Domínguez-Vega, E.; Nouta, J.; Pongracz, T.; de Reijke, T. M.; Wuhrer, M.; Lageveen-Kammeijer, G. S. M. Profiling the Proteoforms of Urinary Prostate-Specific Antigen by Capillary Electrophoresis – Mass Spectrometry. *Journal of Proteomics*. **2021**, *238*. DOI: 10.1016/j.jprot.2021.104148.

(78) Ashwood, C.; Pratt, B.; MacLean, B. X.; Gundry, R. L.; Packer, N. H. Standardization of PGC-LC-MS-Based Glycomics for Sample Specific Glycotyping. *Analyst*. **2019**, *144* (11), 3601–3612. DOI: 10.1039/c9an00486f.

(79) Selman, M. H. J.; Derks, R. J. E.; Bondt, A.; Palmblad, M.; Schoenmaker, B.; Koeleman, C. A. M.; van de Geijn, F. E.; Dolhain, R. J. E. M.; Deelder, A. M.; Wuhrer, M. Fc Specific IgG Glycosylation Profiling by Robust Nano-Reverse Phase HPLC-MS Using a Sheath-Flow

ESI Sprayer Interface. *Journal of Proteomics*. **2012**, *75* (4). DOI: 10.1016/j.jprot.2011.11.003.

(80) Higel, F.; Demelbauer, U.; Seidl, A.; Friess, W.; Sorgel, F. Reversed-Phase Liquid-Chromatographic Mass Spectrometric N-Glycan Analysis of Biopharmaceuticals. *Anal Bioanal Chem*. **2013**, *405* (8), 2481–2493. DOI: 10.1007/s00216-012-6690-3.

(81) Sénard, T.; Gargano, A. F. G.; Falck, D.; de Taeye, S. W.; Rispens, T.; Vidarsson, G.; Wuhrer, M.; Somsen, G. W.; Domínguez-Vega, E. MS-Based Allotype-Specific Analysis of Polyclonal IgG-Fc N-Glycosylation. *Frontiers in Immunology*. **2020**, *11*, 2049. DOI: 10.3389/fimmu.2020.02049.

(82) Bruggink, C.; Wuhrer, M.; Koeleman, C. A. M.; Barreto, V.; Liu, Y.; Pohl, C.; Ingendoh, A.; Hokke, C. H.; Deelder, A. M. Oligosaccharide Analysis by Capillary-Scale High-PH Anion-Exchange Chromatography with on-Line Ion-Trap Mass Spectrometry. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*. **2005**, *829* (1–2). DOI: 10.1016/j.jchromb.2005.10.006.

(83) Madunić, K.; Zhang, T.; Mayboroda, O. A.; Holst, S.; Stavenhagen, K.; Jin, C.; Karlsson, N. G.; Lageveen-Kammeijer, G. S. M.; Wuhrer, M. Colorectal Cancer Cell Lines Show Striking Diversity of Their O-Glycome Reflecting the Cellular Differentiation Phenotype. *Cellular and Molecular Life Sciences*. **2021**, *78* (1), 337–350. DOI: 10.1007/s00018-020-03504-z.

(84) Smith, J.; Millán-Martín, S.; Mittermayr, S.; Hilborne, V.; Davey, G.; Polom, K.; Roviello, F.; Bones, J. 2-Dimensional Ultra-High Performance Liquid Chromatography and DMT-MM Derivatization Paired with Tandem Mass Spectrometry for Comprehensive Serum N-Glycome Characterization. *Analytica Chimica Acta*. **2021**, *1179*, 338840. DOI: 10.1016/j.aca.2021.338840.

(85) Kuhn, R.; Hoffstetter-Kuhn, S. *Capillary Electrophoresis: Principles and Practice*; **1993**. DOI: 10.1007/978-3-642-78058-5.

(86) Kammeijer, G. S. M. *Unravelling the Sugar-Coating of Prostate-Specific Antigen: Method Development and Its Application to Prostate Cancer Research*; Leiden University, **2019**.

(87) Schmitt-Kopplin, P.; Fekete, A. The CE-Way of Thinking: “All Is Relative!” In *Methods in Molecular Biology*; **2016**; Vol. 1483. DOI: 10.1007/978-1-4939-6403-1_1.

(88) Kammeijer, G. S. M.; Jansen, B. C.; Kohler, I.; Heemskerk, A. A. M.; Mayboroda, O. A.; Hensbergen, P. J.; Schappler, J.; Wuhrer, M. Sialic Acid Linkage Differentiation of Glycopeptides Using Capillary Electrophoresis - Electrospray Ionization - Mass Spectrometry. *Scientific Reports*. **2017**, *7* (1). DOI: 10.1038/s41598-017-03838-y.

(89) Rohrer, J. S.; Thayer, J.; Weitzhandler, M.; Avdalovic, N. Analysis of the N-Acetylneuraminic Acid and N-Glycolylneuraminic Acid Contents of Glycoproteins by High-PH Anion-Exchange Chromatography with Pulsed Amperometric Detection (HPAEC/PAD). *Glycobiology*. **1998**, *8* (1). DOI: 10.1093/glycob/8.1.35.

(90) Ruhaak, L. R.; Huhn, C.; Waterreus, W. J.; de Boer, A. R.; Neususs, 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. DOI: 10.1021/ac800630x.

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

(92) Madunić, K.; Wagt, S.; Zhang, T.; Wuhrer, M.; Lageveen-Kammeijer, G. S. M. Dopant-Enriched Nitrogen Gas for Enhanced Electrospray Ionization of Released Glycans in Negative Ion Mode. *Analytical Chemistry*. **2021**, *93* (18). DOI: 10.1021/acs.analchem.1c00023.

(93) Ruhaak, L. R.; Deelder, A. M.; Wuhrer, M. Oligosaccharide Analysis by Graphitized Carbon Liquid Chromatography-Mass Spectrometry. *Analytical and Bioanalytical Chemistry*. **2009**, *394* (1). DOI: 10.1007/s00216-009-2664-5.

(94) Zauner, G.; Deelder, A. M.; Wuhrer, M. Recent Advances in Hydrophilic Interaction Liquid Chromatography (HILIC) for Structural Glycomics. *Electrophoresis*. **2011**, *32* (24). DOI: 10.1002/elps.201100247.

(95) Alpert, A. J. Hydrophilic-Interaction Chromatography for the Separation of Peptides, Nucleic Acids and Other Polar Compounds. *Journal of Chromatography A*. **1990**, *499* (C). DOI: 10.1016/S0021-9673(00)96972-3.

(96) Camperi, J.; Pichon, V.; Delaunay, N. Separation Methods Hyphenated to Mass Spectrometry for the Characterization of the Protein Glycosylation at the Intact Level. *Journal of Pharmaceutical and Biomedical Analysis*. **2020**, *178*. DOI: 10.1016/j.jpba.2019.112921.

(97) Lermyte, F.; Tsybin, Y. O.; O'Connor, P. B.; Loo, J. A. Top or Middle? Up or Down? Toward a Standard Lexicon for Protein Top-Down and Allied Mass Spectrometry Approaches. *Journal of the American Society for Mass Spectrometry*. **2019**, *30* (7). DOI: 10.1007/s13361-019-02201-x.

(98) Yang, Y.; Franc, V.; Heck, A. J. R. Glycoproteomics: A Balance between High-Throughput and In-Depth Analysis. *Trends in Biotechnology*. **2017**, *35* (7), 598–609. DOI: 10.1016/j.tibtech.2017.04.010.

(99) Lageveen-Kammeijer, G. S. M.; Kuster, B.; Reusch, D.; Wuhrer, M. High Sensitivity Glycomics in Biomedicine. *Mass Spectrometry Reviews*. **2022**, *41* (6). DOI: 10.1002/mas.21730.

(100) Zhu, Z.; Desaire, H. Carbohydrates on Proteins: Site-Specific Glycosylation Analysis by Mass Spectrometry. *Annual Review of Analytical Chemistry*. **2015**, *8*. DOI: 10.1146/annurev-anchem-071114-040240.

(101) Ruhaak, L. R.; Xu, G.; Li, Q.; Goonatilleke, E.; Lebrilla, C. B. Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chemical Reviews*. **2018**, *118* (17). DOI: 10.1021/acs.chemrev.7b00732.

(102) Varki, A. et al. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology*. **2015**, *25* (12), 1323–1324. DOI: 10.1093/glycob/cwv091.