Unravelling the sugar-coating of prostate-specific antigen : method development and its application to prostate cancer research

Kammeijer, G.S.M.

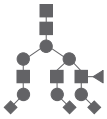
6

DISCUSSION AND FUTURE PERSPECTIVES

DISCUSSION AND FUTURE PERSPECTIVES

The aim of this thesis was to explore the analysis of glycans and glycopeptides using the powerful analytical platform CE-ESI-MS (capillary electrophoresis - electrospray ionization - mass spectrometry) employing a porous nano-sprayer for the MS coupling. In addition, the developed platform was used to study in-depth the glycosylation of prostate-specific antigen (PSA). This platform was included in a PSA Glycomics Assay (PGA) that captures PSA from urine in order to study the glycosylation profiles of PSA from patient samples and evaluate their potential as a biomarker for patient stratification in prostate cancer (PCa).

SENSITIVITY ENHANCEMENT WITH CE-ESI-MS



Even though CE-ESI-MS is known for its limited sample injection it still is able to achieve a high sensitivity, especially when using a sheathless electrospray ionization interface. In this thesis, the sensitivity was further pushed by implementing a dopant enriched nitrogen (DEN)-gas. Significantly higher sensitivities and an improved repeatability were observed for glycopeptide (**Chapter 3**) and glycan analysis (**Chapter 4**). The implementation of the DEN-gas resulted in a 25-fold increase of sensitivity compared to conventional sheathless CE-ESI-MS platform, with a glycopeptide amount of 0.2 amol injected representing the lower limit of detection. Interestingly, a similar study was performed in 2013,[177] where a limit of detection (LOD) of 20 amol was determined for immunoglobulin G (IgG) glycopeptides on a conventional sheathless CE-ESI-MS platform, revealing a 4-fold increase in sensitivity between the two conventional CE-ESI-MS platforms (present (no DEN-gas) *versus* the study in 2013). It should be noted that a higher sample injection was used in 2013 with a capillary volume of 37% compared to 7% in the current study. The former most likely compromised the separation power resulting in a lower signal-to-noise ratio and a decrease in sensitivity. Thereby, an overall 100-fold increase in sensitivity could be observed when the DEN-gas was implemented compared to the sheathless CE-ESI-MS platform in 2013.

One may speculate with respect to the mechanisms underlying the observed enhancement in sensitivity when various dopants are introduced during electrospray ionization. Two electrospray ionization mechanisms that may be involved are the **charge residue model (CRM)**, which was already proposed in 1968 by Dole *et al.* and the **ion evaporation mechanism (IEM)** first described by Iribarne and Thomson in 1975 (**Figure 6.1**).[247] Both models describe possible mechanisms for the formation of gas-phase ions in electrospray ionization. The CRM starts with the assumption that evaporation of the solvent decreases the size of a droplet

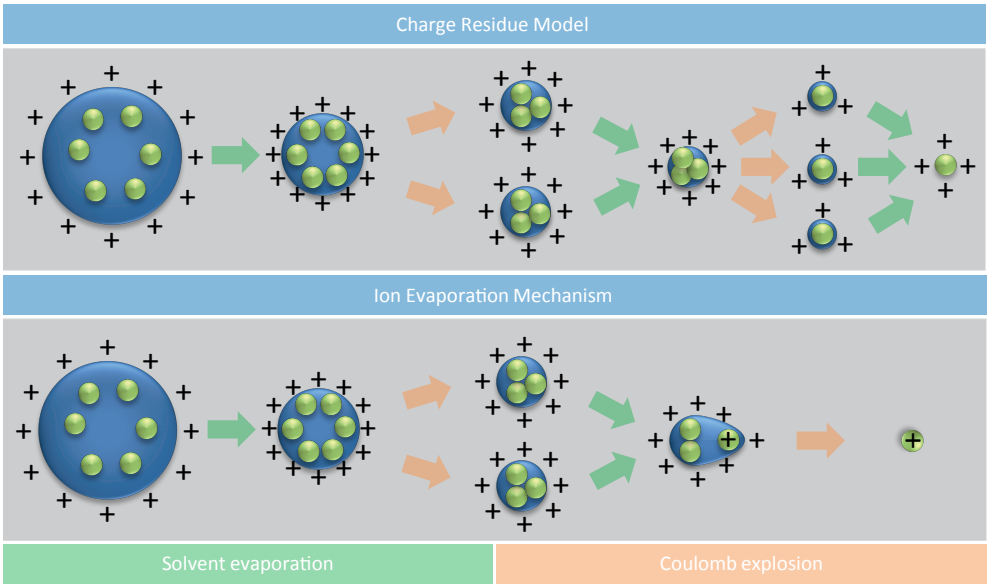
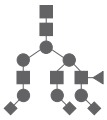


Figure 6.1: A schematic illustration of two different models of the formation of gas-phase ions from a solvent droplet; charge residue model (CRM) and ion evaporation mechanism (IEM). The main difference between the two different models is that gas-phase ions in CRM are generated by evaporation of the solvent surrounding the analyte and charges of the solvent are being transferred on the analyte. IEM suggests that the gas-phase ions are formed through evaporation followed by desorption.

until its **Rayleigh limit** is reached (maximum level of tension a droplet can handle), leading to the division of the droplet into smaller droplets (**Coulomb explosion**). This process will repeat itself until final droplets are formed, containing a single analyte. When the last solvent surrounding the analyte evaporates, the charges of the droplet are being transferred onto the analyte resulting in a charged residue.[248] As reported in the literature, this process seems to apply most likely to large molecules.[249] According to the IEM model, the gas-phase ions are considered to be formed due to evaporation and desorption. The starting point is assumed to be the same as for CRM, with the evaporation of the solvent and the division of the droplet into smaller ones; this process will repeat itself until analyte desorption (ejection of solvated ions from the droplet surface) takes place, resulting in the formation of gas-phase ions. This mechanism seems to be most applicable for smaller molecules. The involvement of an IEM-mechanism in the increased sensitivity (when using a DEN-gas) is supported by a study by Nguyen and Fenn in 2007, where the effect of solvent vapors added to the sheath gas was studied and how this influenced the ion formation.[191] By increasing the solvent portion in the gas phase, a better desolvation of ions was observed when the vapor contained polar organic solvents such as methanol or butanol compared to a solvent vapor that only contained water or the non-polar solvent cyclohexane. Moreover, researchers suggest that the difference in gas-phase basicity and proton affinity between MeCN and water could be a crucial factor for the variety in ionization efficiencies and charge state distribution and that this is related

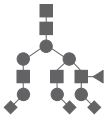
to the CRM mechanism.[211] Both in **Chapters 3** and **4** an increase in multiple protonated ions was found, and it is thought that this is due to the adsorption of solvent vapor molecules to the surface of the droplet, leading to "hot spots" that aid to the ejection of the analyte from the droplet.[191] As presented in **Chapter 3**, the use of different organics (isopropanol (IPA), ethanol (EtOH) and acetonitrile (MeCN)) resulted in a variation in the relative abundances of glycopeptides. Interestingly, IPA and EtOH revealed similar relative peak areas and seemed to favor the less complex analytes with smaller glycan portions such as G0F (H3N4F1) while MeCN seemed to favor more complex types such as G2F (H5N4F1) and G2FS (H5N4F1S1). When comparing the charge state distributions of G0F the most abundant charge state observed with IPA and EtOH as a dopant is the doubly charged species while for MeCN this is the triply charged state. A similar difference is observed for G2FS, were the triply charged species are barely present when IPA or EtOH are used as dopants while they are the most abundant form for MeCN, followed by the quadruply charged state with the doubly charged state is almost non-existing. As previously mentioned, the gas-phase basicity and proton affinity could play a role in this phenomenon as the gas-phase basicity is 660 kJ/mol and 748 kJ/mol for water and MeCN, respectively. The proton affinity is 691 kJ/mol for water and 779 kJ/mol for MeCN, respectively.[250] However, the gas-phase basicity and proton affinities of all dopants used in this study are not that far apart from each other, as EtOH has a gas-phase basicity and proton affinity of 746 kJ/mol and 776 kJ/mol, while this is 763 kJ/mol and 793 kJ/mol for IPA,[250] respectively. Most likely the gas-phase basicity and proton affinity play an important role for the increased sensitivity, however, it does not explain the ionization difference that is observed between the dopants. One of the differences that can be found between the solutions is the fact that MeCN is a weak base in solution (weaker than water), while this is not the case for EtOH and IPA. This implies that MeCN is unlikely to take up protons, and during condensation of MeCN into the droplets, the property of the droplet will change to an overall lower basicity of the solvents, which may promote supercharging of the analyte. Moreover, based on the abovementioned facts, it is plausible that both IEM and CRM are involved in signal modulation and enhancement. However, dependent on the dopant that is being used a different mechanism can be involved. Namely, it could be that IPA and EtOH are more involved in the IEM mechanisms while MeCN seems to be involved in the CRM model. Since the increased sensitivity could be due to a better desorption of the analytes (creating hot-spots of desorption by dopants condensation to the electrospray microdroplets) and, therefore, the IEM model will be most likely. Moreover, EtOH and IPA seem to have a preference for smaller molecules that is in agreement with the IEM model. In the case of MeCN, supercharging is being observed and this is most likely related to the CRM model as, according to this model, charges from the droplet can be transferred onto the analyte resulting in higher charge states, and this dopant favors larger molecules. However, how the ionization exactly takes place remains unclear and further research is needed to better understand this process. Therefore, it is suggested to investigate a range



of less polar dopants (such as pentane or hexane) complementing the abovementioned dopants, in order to confirm that the less polar solvents indeed do not aid in the improved ionization efficiency. Moreover, the addition of volatile acids, such as acetic acid and formic acid, in minor concentrations, should be investigated as well as the influence of the pH of the solution in relation to the observed ESI response. Another interesting study would be the effect of adding organics to the background electrolyte (BGE) and not to the nebulizer gas; however, this would inevitably also influence the separation of the molecules. Further studies should not only focus on the effects of DEN-gas on glycopeptide ionization, but also consider other analyte groups such as small molecules (metabolites), medium-sized analytes molecules (peptides and various modified peptides) and large molecules (intact proteins). By covering a vast range of physicochemical properties of different analytes, it will be possible to gain a better insight into which molecular properties are involved in the sensitivity enhancement upon the use of a DEN-gas. During this thesis, the effect of the DEN-gas was further examined by investigating the effect on released *N*-glycans carrying a cationic label (**Chapter 4**) next to the effect of DEN-gas on glycopeptides. For the labeled *N*-glycans similar results were found to those described in **Chapter 3** for PSA glycodipeptides, with an overall improvement of 3-fold in signal-to-noise ratio as well as a 2-fold increase in peak area. Obtaining similar results for PSA glycopeptides and labeled *N*-glycans was rather expected as the backbone of both analytical species was rather small and most probably only played a small role in the observed enhancement.

GLYCAN AND GLYCOPEPTIDE ANALYSIS WITH CE-ESI-MS

The CE-ESI-MS proved to be an excellent technique for high separation efficiencies, especially beneficial for the separation of isomers with differently linked sialic acids (**Chapter 2**). While other platforms often require extra sample treatment for glycopeptides, such as derivatization or complex analytical setups, this is not the case for CE-ESI-MS. However, the analysis of released *N*-glycans in normal mode CE (detection at the cathodic end of the capillary) did benefit from additional derivatization steps, namely, by reducing the charge heterogeneity of *N*-glycans, mainly caused by the presence of sialic acids (from zero to up to four). By derivatizing the sialylated species, these *N*-glycans were made neutral and then, to enable separation in CE, a positive charge was added allowing analysis in normal mode. This included an additional step where the reducing end was labeled with a permanent cation, Girard's reagent P (GirP). This workflow enabled the identification of 167 *N*-glycoforms with CE-ESI-MS, including differentiation between differently linked sialic acids. In addition, we got insights into the electrophoretic mobility of the *N*-glycans (**Chapter 4**). Namely, bisected structures have a rather small hydrodynamic volume as compared to non-bisected structures of similar molecular mass and, therefore, a higher electrophoretic mobility. This is illustrated by comparing the migration times (t_m) of H3N5 (1653.638 Da) and H4N4 (1612.611 Da) where a difference of 18 s was observed with the bisected structure migrating earlier (35.5 min *versus* 35.8 min) despite the higher molecular mass, which is in agreement with literature.[217] This phenomenon might be useful for the differentiation between isomeric *N*-glycans with antenna-branching and bisecting GlcNAc. A similar difference in electrophoretic mobility can be observed for high mannose types structures, *e.g.,* H8N2 (t_m: 36.1 min, 1854.664 Da), which have a rather high electrophoretic mobility compared to complex structures, *e.g.,* H5N4 (t_m: 36.2, 1774.664 Da) while having a higher molecular weight. However, no direct correlation could be found between the hydrodynamic volume of differently linked sialic acids of mono-sialylated species and their electrophoretic mobility, *e.g.,* H5N4Am1 (α2,3 sialylated, 2064.775 Da) and H5N4E1 (α2,6 sialylated, 2093.791 Da) both migrated at 36.8 min. While the di-antennary species did result in slight different mobilities, *e.g.,* H5N4F1Am2 (*m/z* 2500.945 Da), H5N4F1Am1E1 (2529.960 Da), H5N4F1E2 (2558.975 Da) migrated at 37.6 min, 37.4 min and 37.4 min, respectively. Notably, a tri-antennary *N*-glycan with two *N*-acetylneuraminic acids (H6N5Am1E1, 2749.034 Da) resulted in a later migration (37.8 min) than a tetra-antennary *N*-glycan with one *N*-acetylneuraminic acid (H7N6E1, 2824.055 Da) detected at 37.7 min, suggesting that the presence of two *N*-acetylneuraminic acids has a larger impact on the hydrodynamic volume than the addition of an antenna (galactose + *N*-acetylglucosamine).



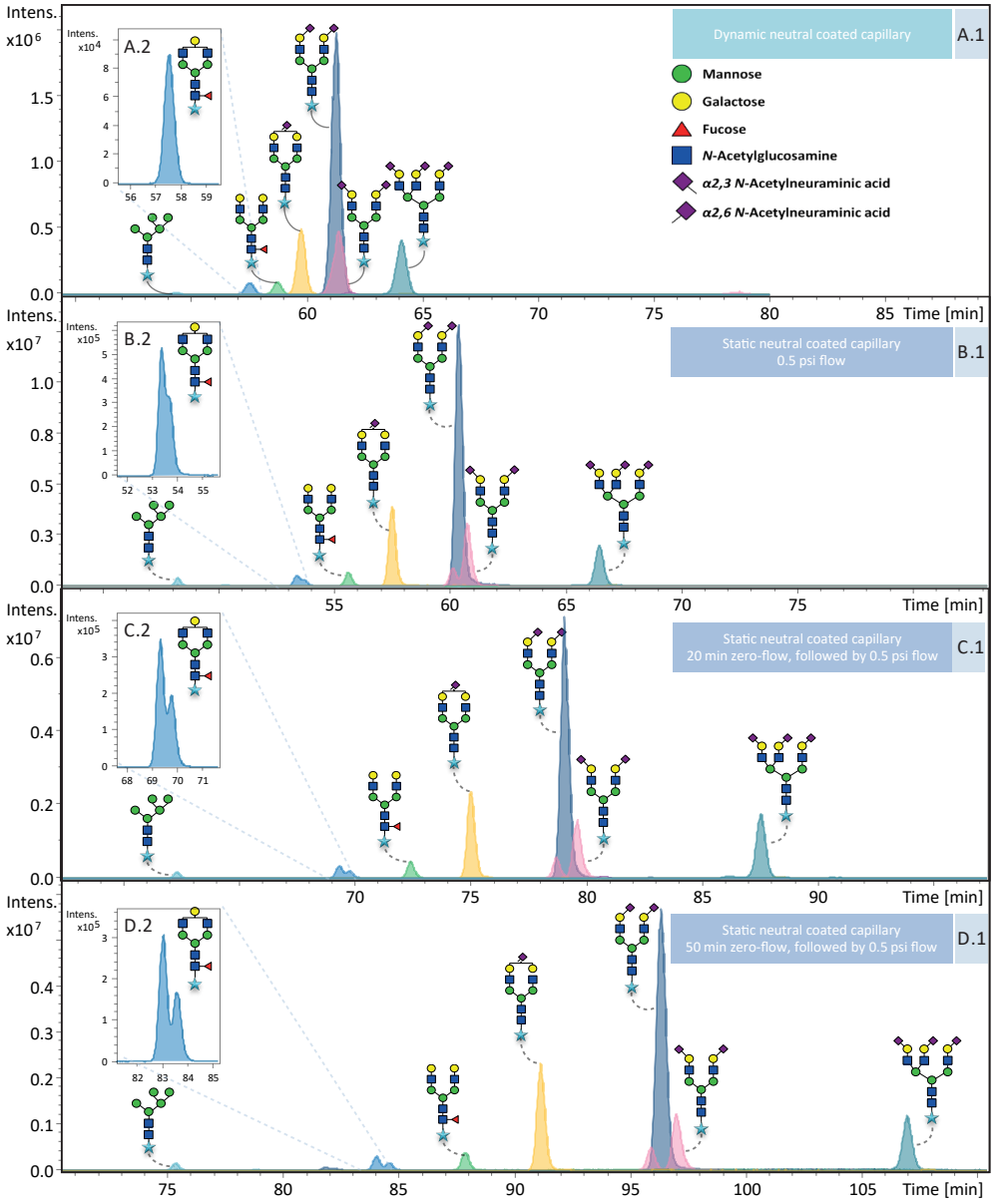


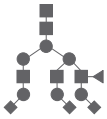
Figure 6.2: Extracted ion electropherograms of total plasma *N*-glycome after derivatization (ethyl-esterified amidation) and labeling with GirP. (A.1) Separation observed for TPNG glycans using a dynamic neutral coated bare fused silica (BFS) capillary, **(B.1-C.1)** Separation observed for TPNG glycans using a static neutral coated BFS capillary, with either **(B)** a constant 0.5 psi flow, **(C)** 0.5 psi flow after 20 min or **(D)** 0.5 psi flow after 50 min. **(A.2-D.2)** Insets show the extracted ion electropherogram of *N*-glycan species H4N4F1. All analyses were performed with CE-ESI-MS using the DEN-gas. GirP illustrates the label, Girard's reagent P, attached to the glycans. The following settings were used: The background electrolyte was 10% HAc, the temperature of capillary coolant was set at 25˚C, 20kV was applied and injection was performed with 1 psi for 60 sec (9 nL), sample was diluted in 100 mM ammonium acetate pH 4.0.

However, to further explore structural differentiation in relation to hydrodynamic volume, higher separation efficiency is needed. The use of a neutrally coated sheathless capillary could provide the desired resolution. Here, as the electro-osmotic flow (EOF) is very low, the mobility of the analytes will be mainly determined by their own electrophoretic mobility. Unfortunately, to enable detection by MS, a constant flow should be created to generate an electrospray, hampering the separation power of a neutral capillary. This flow can be created by applying a mechanical pressure, which will, however, disturb the advantageous planar flow created by the EOF. Instead a parabolic flow profile could be created resulting in broader peaks and a loss in resolution. While a dynamic neutral coating (Ultratrol) was used in **Chapter 4**, this did not result in an efficient suppression of the EOF as an electrospray was obtained while no mechanical pressure was applied, suggesting that the capillary was not completely coated with the dynamic neutral coating or that the coating is actually not neutral. Therefore, a static coating was explored which resulted in sufficient suppression of the EOF, as no electrospray was observed when no mechanical pressure was applied. To initiate electrospray ionization a continuous pressure of 0.5 psi was applied (**Figure 6.2.A.1**). By applying a zero-flow principle (no flow) for a certain period right after sample injection, a better separation could be achieved. Just recently, this process has been explored and preliminary data revealed promising results with extraordinary resolution, allowing the resolution of isomeric species (**Figure 6.2.B.1** till **D.1**). As illustrated in the insets **A.2** till **D.2** of **Figure 6.2** for glycan species H4N4F1, isomer separation improved by applying the zero-flow principle for a longer period. Based on literature, the first migrating species is most likely the species with a galactosylated α6 arm, while the second migrating peak has an occupied α3 arm.[251] Mechref and coworkers further investigated the separation by computer modeling.[251] They demonstrated that the glycan with an occupied α6 arm is more compact than its isomeric species, due to the closer proximity of the core-fucose and the galactose (~6.7 Å distance) when compared to the galactose on the α3 arm (~14.1 Å).[251] Notably, the galactosylated α3 did not only show a slower mobility in CE, but also less retention on porous graphitized carbon liquid chromatography (PGC-LC)-ESI-MS(/MS).[251] However, further research is needed to structurally characterize the various isomeric species. For this purpose, endo- and exoglycosidases, together with well-characterized glycan standards, could be used. In addition, as the zero-flow process comes with rather long analysis time, further optimization is necessary and the use of different BGE, the addition of organic modifiers, but also the applied voltage, capillary temperature, pH and composition of the leading electrolyte should be optimized with respect to effects on separation and analysis time.[237,238] Another aspect to be investigated is the use of labels containing more than one charge, either permanently charged or chargeable under specific CE conditions, and see how this influences separation and the MS detection. The ion desolvation during the ESI could be further improved by investigating more hydrophobic labels.[234] Furthermore, studies are needed that compare the developed platform to other well-established assays

involving labeling with negatively charged (or chargeable) labels on reversed mode CE-ESI-MS platforms with negative mode MS detection such as the study by Snyder *et al.* where the label APTS (8-aminopyrene-1,3,6-trisulfonic acid) was used.[216] These studies should provide information on which labels are most suitable for isomeric separation, to which extent and how well they can be detected (sensitivity). For this, it should be taken into account that an overall lower sensitivity is expected when labeled *N*-glycans are detected in negative ionization mode MS, as compared to positive ionization mode,[239] which would be the case for APTS labeled glycans. However, in comparison to the "zero-flow" analysis, the APTS labeled glycans appear to migrate fast and in a rather small separation window with an excellent resolution that can be seen as a major advantage.[252] Moreover, no sialic acid derivatization is needed when samples are measured in negative ionization mode MS. Additionally, the negative charge of the sialic acids leads to faster migration and an increase of the separation window. APTS-labeled glycans are normally analyzed on CE-laser-induced fluorescence (LIF) or CE-ultraviolet (UV) platforms, which tend to be extremely repeatable at rather low-costs. One of the key attributes to its repeatability is the detectors that are being used and in many analytical settings LIF or UV detection will suffice. Only when in-depth information is required, for example when new peaks are found in the electropherogram or when there is an expectancy of overlapping analytes, the CE has to be coupled to MS to identify the analytes under peaks observed in the electropherogram. Therefore, it is suggested to study whether the cationic label GirP is also detectable with a UV or a LIF detector allowing for simplified analyses without MS detection.

PERSONALIZED MEDICINE

Nowadays, more and more research is devoted to **personalized medicine**, which is, as stated by Nature,[253] "a *therapeutic approach involving the use of an individual's genetic and epigenetic information to tailor drug therapy or preventive care.*" This means that every individual is treated taking into account his/her health, lifestyle and the disease state, leading to a personalized treatment, which is expected to optimize effectiveness and minimize side effects, while providing an improvement in the quality of life. This is especially relevant when diseases are found in an early stage and a direct tailored treatment can be initiated. Furthermore, personalized medicine should lead to lower treatment costs due to a faster recovery of the patients or to the inhibition of the disease progression. This principle is gaining importance in the health care sector and receives increased attention from governments and funding agencies. For example, the Dutch government wrote in the coalition agreement of 2017 - 2021 that personalized medicine is seen as one of the innovations in relation to the prescription of pharmaceuticals for patients.[254] Personalized medicine is increasingly making use of molecular signatures, in which glycomics and glycoproteomics are emerging fields. Here, glycosylation profiles could provide an indication of the health of the patient.[87,97,255,256] Though preliminary studies show promising results for using glycoprotein signatures in early diagnosis, patient stratification and treatment monitoring, more studies are still required to determine their significance.[87,94,256]



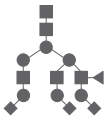
Specific molecular features that can be correlated to a certain disease are known as clinical **biomarkers**, described by the National Cancer Institute as "*a biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease. A biomarker may be used to see how well the body responds to a treatment for a disease or condition. Also called molecular marker and signature molecule.*"[257] Therefore, a biomarker can provide insight into whether a patient has a certain disease, and if so, how to proceed within the clinical pathway. It needs to be noted that a biomarker can consist of a single molecule or a set of molecules that can provide a certain signature for the disease. Additionally, the obtained information contributes to the knowledge on the molecular pathway and potential irregularities that are related to a certain disease, enabling patient risk stratification, *e.g.,* in early screening, or the differentiation between disease states such as indolent and aggressive tumors.[258]

BIOMARKER DISCOVERY

It is remarkable that on a yearly basis many potential clinical biomarkers are identified, while only a minor fraction will eventually find its way into the clinics.[259] The process of translating a potential biomarker into the clinic is slow and often takes many years, making it time-consuming and costly while there is a need for instant new biomarkers. The slow progress and the very limited translation is often a result of poorly designed experiments in the discovery phase, due to a lack of communication between the different collaborators, also known as **stakeholders**, such as the clinicians, industry, academic researchers and clinical laboratory professionals, even though they may all strive for improving the diagnosis and prognosis of a disease.

The academic researcher may often be focused on finding new biomarkers and overlook which criteria apply for the development of a new clinical test or critical steps that largely influence the robustness of a developed assay (*e.g.,* sample collection, storage conditions), leading to irreproducible results. While the clinical laboratory professional might not be fully aware of the fact that not all critical steps can be thoroughly tested and taken into account when a new biomarker is still in the discovery phase (and, therefore, its potential has not yet been validated). It should be noted that, aside from more specific biomarkers, there is also a need for a better translation from the discovery phase into the validation phase. For example, it should be taken into account that most analytical platforms that are used in the biomarker discovery phase are chosen for their ability to analyze and characterize the sample to a considerable depth to cover as many features as possible. Therefore, different requirements are often set for the desired analytical platform in the biomarker discovery phase compared to the clinical setting, leading to platforms that are most often not suitable in a clinical setting. Resulting in the fact that the translation should not only be focused on validating the outcomes, but also on the development of more simplistic assays that allow high-throughput and robustness. Moreover, it would be beneficial if the different stakeholders jointly identify at an early stage the crucial factors with respect to sample collection, transportation and handling during pre-, mid- and post-processing, also known as the **pre-analytics**.[258,260] Examples of questions to be addressed are: *How do the samples have to be collected? Is the sample container crucial? Could the matrix of the sample bias the results? Is there a need for a stabilization agent? Does the sample require direct storage or is it required to cool the sample down to a certain temperature prior to storage? At which temperature does the sample need to be stored? Do several freeze-and-thaw cycles change the sample composition?* All these factors can be critical for the outcome of a biomarker assay as well as influence the reproducibility and reliability of the assay. Therefore, prior to testing an established assay within a small cohort, the whole workflow should be critically evaluated and an inventory of crucial steps should be made. As a result, the needs and critical points are better understood and repetition of experiments can be avoided when transferring from

the discovery phase to the validation phase, eventually saving time and resources, providing more reliable results and giving more confidence in an actual biomarker that will find its way into the clinics. Moreover, the **analytical validity** should also be evaluated, which takes into account the technical aspects of an assay.[258,260] Here, the robustness of an assay is determined, such as its accuracy, repeatability, but more importantly its reproducibility, which provides information about how well the assay performs when performed by different analysts, with other instruments and at different laboratories. By testing pre-analytical factors and clinical validity, the outcomes become more reliable since findings can be reproduced. Another crucial point of a successful biomarker is its **clinical validity**.[258,260] The biomarker should provide evidence that it is able to distinguish patients with or without a disease. In the case of PCa, this can be the distinction between indolent and aggressive PCa or between benign prostate hyperplasia (BPH) and PCa patients. Whenever a biomarker seems to be clinically valid it should be reproduced with an independent sample set; whether the original researchers should perform this, is still under debate.[258,259] Biomarkers that do not provide any advantage on assays that are currently used in the clinical pathway do not show their **clinical utility** and in the end they will not be utilized.[258,260] Biomarkers are expected to address "*any missing or inadequately performing component of a clinical pathway*", which is defined by the Test Evaluation Group of the European Federation of Clinical Chemistry and Laboratory Medicine as an **unmet clinical need**.[261] This federation proposed a 4-step process that helps to identify these unmet clinical needs in a clinical pathway.[260,261] The **first step** is scrutinizing literature, critically evaluating described biomarkers, identifying the current practice, the gap in the clinical pathway and what the desired outcomes are if these gaps are filled. The **second step** investigates whether there is already an existing and satisfying solution for this gap. If yes, there is no unmet clinical need, if not, a study is needed to identify a possible solution. The **third step** includes the question if a potential biomarker would contribute to the solution and what its impact would be and where in the clinical pathway it should be placed. For example, would it be included as an additional, a replacement or a triage test (determining which patient should be treated immediately). What would be the advantage of including this test, what are the potential harms and will the advantages expectedly outweigh these harms. The **fourth** and **final step** considers if the proposed solution would be practically feasible, not only in terms of technical feasibility, but also with regard to performance and standardization of the assay, as well as capital investment. The commercial feasibility is also taken into account by defining its cost-effectiveness, *e.g.,* how well the new test performs compared to the current process and are the costs worth the improvements. Another feasibility evaluation is made on what is needed on an organizational level: will the improvement be acknowledged by the patients and all stakeholders and, if not, what is needed to change this? What facilities are required? Is there a special training needed for the staff? The accessibility should also be taken into account, *e.g.,* will this be a test that can be performed in every hospital or are there specialized laboratories needed and, if so, what would be the availability and accessibility in this case?



PSA GLYCOSYLATION AS POTENTIAL BIOMARKER

The last section of this thesis is focused on the **application** of PSA glycosylation analysis and its evaluation as a PCa biomarker. **Chapter 5** describes the development of a PSA Glycomics Assay, which allows the capture of intact PSA from patients' urine followed by tryptic digestion and the analysis with the optimized CE-ESI-MS platform (**Chapters 2** and **3**).

In the case of PCa, there are two unmet clinical needs: the need for a better marker in the early diagnosis of PCa and, most importantly, the differentiation between indolent from aggressive tumors. Early diagnosis of the disease would lead to early treatment and less chance for disease progression. While early screening tests seem to be a crucial step in the clinical pathway, the current PSA test lacks specificity, leading to many false positives resulting in unnecessary investigations.[262] Concerning the differentiation between indolent and aggressive disease, patients with low-grade PCa (Gleason score ≤ 6) often do not require immediate treatment while those with aggressive PCa (Gleason score ≥ 7) do.[24] However, as the current PSA test is unable to identify the aggressiveness of the cancer, additional tests are needed such as biopsies, which are invasive and can be harmful for the patient (pain, anxiety and infections). Biopsies can provide more information whether a tumor is present and how aggressive it is, however, the outcome is operator dependent and this only reflects the status of the biopsy tissue. Namely, the cancer stage of the patient may not be accurately reflected as it does not always represent the status of the entire prostate gland. In this respect, if a positive biopsy is found it can be confirmed if a patient has PCa, however, the most aggressive form found in the biopsy can only define the disease status. This could result in undertreatment when (more) malignant tumors were missed during biopsy. This indicates that there is still a gap in accurate identification of PCa and its aggressiveness. Therefore, a potential biomarker should be evaluated on its specificity and ability to provide more in-depth information regarding the disease status of the patient. Once an elevated PSA-concentration has been detected, a biopsy will be taken and a second assay is needed to provide more confidence that a biopsy is truly negative and there is indeed no indication for the presence of an aggressive PCa. When this assay does not provide any indication no further investigations will be needed other than repeat testing of PSA as well as performing the newly developed secondary assay (**step 1**). Even though several attempts have been made to overcome this unmet need, none of them have provided enough evidence of their advantages over the use of the current PSA test (**step 2**).[78,80] However, the glycan modifications of PSA can be a useful biomarker. Namely, based on preliminary results, the glycosylation profile of PSA, especially alterations in the molecular features (sialylation and fucosylation)[3,35,99,101-103] could improve the PCa diagnosis leading to better patient stratification. The minimum clinical performance that is required for a new biomarker is at least a better specificity compared to the gold standard (PSA test in combination with a biopsy). This would be achieved if indolent PCa can be distinguished from aggressive PCa with a lower number of false positives (PSA

test) and false negatives (biopsy) (**step 3**). Finally, the feasibility of the identified biomarker should be studied (**step 4**), however, this is dependent on how the biomarker is defined. Ideally, a platform will be developed that will be largely operator independent, providing a certain value that advises the clinician how to proceed further, *e.g.,* is there a need for more in-depth examination or would active surveillance be a better option for this patient and if so, in which time period they should be examined again. Currently, patients under active surveillance will have a PSA test every three months and a biopsy after one year, making this a very bothering process for the patients. Ideally, the value from the test should distinguish patients that need examination *e.g.,* every three months from those that need an examination every two years or even five years, that would significantly improve the quality of life for a patient.

Additionally, it is worth noting that the biopsy is acknowledged as the gold standard for PCa diagnosis, however, this is a rather biased gold standard as this is mainly based on histology which is operator dependent and, moreover, a tumor could be missed during biopsy leading to misdiagnosis. Therefore, a suitable clinical endpoint would be *e.g.,* a 10-year long follow-up and see if the disease progresses over time. However, this would still be biopsy dependent. Another option would be radical prostatectomy as this could provide more accurate information about the current status of the disease as the whole prostate can be examined. However, this procedure will lead to information loss in the case of indolent PCa, as it will be uncertain if the disease would have progressed over time. Taking this into account, none of the abovementioned proposals seems to be a perfect gold standard to evaluate a new biomarker which makes it rather complicated to find biomarkers that perform better than a biopsy.

PSA GLYCOMICS ASSAY

The PSA Glycomics Assay developed at our department (**Chapter 5**) was used to study the PSA glycosylation profile of captured PSA from patients' urine and showed that an in-depth relative quantitation could be performed. While the PGA showed excellent repeatability over different days, no direct correlations could be made between the clinical and diagnostic value of PSA glycosylation features in urine. It was hypothesized that this might be due to the small sample size, however, a larger cohort did not result in any new findings (data not shown), suggesting that the urinary PSA glycosylation profile does not have any biomarker potential in early assessment of PCa. The epithelial cells secrete PSA, and in the case of PCa these cells are altered, which could lead to alterations in PSA as well as in its glycosylation (which is initiated in the lumen of the endoplasmic reticulum).[97] It is hypothesized that, in PCa, the PSA that enters the circulation is more derived from altered cells, because the cells are more misaligned and more PSA can leak into the circulation. On the other hand, urine contains already high amounts of PSA under normal conditions, and might not contain a high amount

of the altered PSA. Therefore, it would be of special interest to compare the urinary to the serum PSA glycosylation of patients, to test whether differences can be observed between the two biofluids. The urinary PSA glycosylation profile would provide an unbiased (unaffected by PCa) profile, and this could be used as a baseline to test whether the PSA glycosylation appears changed in the serum. This is especially interesting as the glycosylation between individuals can be rather variable while each patient has its own consistent glycosylation fingerprint, which would make such an assay perfect for personalized medicine.[106]

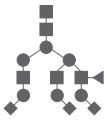
However, many different factors should be taken into account, such as: *does the matrix influence the glycosylation? Does the concentration difference play a role and bias the glycosylation analysis? How stable is PSA in both biofluids? Do different PSA subforms contain different glycosylation features and, if so, can this influence the overall PSA glycosylation profile?* Therefore, more research is needed to study in-depth the biosynthetic pathway of PSA and its *N*-linked glycosylation. Because little is known about when the glycoprotein enters the biofluid, a better insight in this process would provide knowledge on why unglycosylated PSA seems to be more abundant in urine while (differently) nicked PSA subforms seem to be more present in seminal fluid (data not shown). Knowing this background would also help to further explore the possibilities for treatment of the disease. While not covered in this thesis, the intact analysis of PSA could provide the desired knowledge, by identifying the presence and abundance of these different subforms. Different analytical platforms are available which could provide information on different levels, such as based on rough mass differences (SDS-PAGE gel or two-dimensional polyacrylamide gel electrophoresis),[56,244] or charge distribution by anion-exchange chromatography or CE-ESI-MS.[56,263]

Preliminary studies with CE-ESI-MS on the intact analysis of PSA captured from patients' urine showed that, next to the conventional PSA, many different subforms could be identified in urinary PSA (data not shown). Subforms such as inactive PSA (iPSA) and benign PSA (bPSA) could be found, and bPSA was found in different varieties (with one or more internal cleavages and with non, one or two amino acids losses), similar subforms were previously identified in PSA derived from seminal plasma.[56,264] Furthermore, no complexed PSA (cPSA) was found, while bPSA and conventional PSA were found glycosylated as well as unglycosylated with H5N4F1S2 being the most abundant glycan for both forms. These results reveal the added value to study PSA as an intact glycoprotein, next to the analysis on a glycopeptide level. Additional research should focus on further development of the analytical platform by pushing its sensitivity, resolution power and complete assignment of the detected analytes. Additionally, other glycoproteins correlated with PCa can be studied for their biomarker potential such as prostatic acid phosphatase. This protein is known as the predecessor of PSA as it has been used as a serum marker of PCa as it also seems to be elevated in prostate related diseases.[265,266] Another interesting study would be the analysis of various cell lines

and its glycosylation, which hopefully could provide more knowledge about the disease. Until now, most research has been performed on the cell line LNCaP (supraclavicular lymph node, androgen-sensitive human carcinoma of the prostate), however, this cell line seems to have a complete different PSA glycosylation distribution than that observed in urine or plasma, with less sialylation, a higher amount of fucosylation and antennaes.[34,103,267] Therefore, it is suggested to further explore different cell lines that are related to PCa.

BIOMARKER DISCOVERY WITH CE-ESI-MS

As stated previously, information regarding disease-specific changes could provide in-depth knowledge about alterations of the physiological properties of a glycoprotein. Even minor changes as well as low abundant glycosylation features could potentially lead to the desired glycan biomarkers. In this thesis, implementations were made to the CE-ESI-MS platform to further improve the sensitivity in combination with a sheathless interface, and to use these improvements for biomarker discovery in the field of PCa. The sensitivity of the CE-ESI-MS platform has been assessed and compared with other well-established analytical platforms such as nanoLC-ESI-MS (**Chapter 3**) and MALDI-TOF-MS (**Chapter 4**), and has shown to outperform these platforms. Here, the highly sensitive CE-ESI-MS method allowed identifying slight changes and minor abundant glycosylation features that would be missed with other analytical platforms.



It should be noted that the use of CE-ESI-MS(/MS) in a clinical environment is, for now, not envisioned and should be mainly adapted for biomarker discovery, because the analytical platform still has its limitations (no high-throughput and limited sample loadability) and might be too complex to be used in a clinical setting (*e.g.,* operation of the system, repeatability and data complexity). Therefore, the current vision would be to use the obtained knowledge from the in-depth measurements in order to define simpler assays for the clinical setting that focus on the specific features that were found in these studies, allowing high-throughput and more robustness (*e.g.,* lectin arrays, immunoaffinity, aptamer binding assays).[255,256,268]

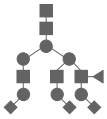
If the current drawbacks of CE-ESI-MS could be resolved, it may have a future in a clinical setting. For example, sample loadability is currently seen as a major disadvantage, as, normally, no more than 2% of the capillary should be filled to avoid loss of resolution,[269] and this often means a sample injection of 9 nL up to a maximum of 70 nL, which is only a minor fraction (0.5% up to 3.5%) of the total volume that is required in the sample vial (2 µL). This is in major contrast to other analytical platforms where larger volumes can be applied (for nanoLC-ESI-MS the volume is up to 10 µL, while for MALDI-TOF-MS it is usually 1 - 2 µL), containing a large part of the sample. Therefore, further developments are needed that enable higher injection volumes in order to enhance the overall sensitivity of CE-ESI-MS workflows. Several studies have already shown promising results;[270] *e.g.,* implementation

of a monolithic column allowing an on-line pre-concentration step of the sample prior to analysis,[200,271,272] the development of new sheathflow and sheathless interfaces[273-275] as well as the use of immunoaffinity.[276] Additionally, the long analysis time could also be seen as another drawback of the current CE-ESI-MS platform and further investigations could be focused on the use of multi-segment injection. During this process several samples (*e.g.,* $N = 5$) can be injected with a BGE plug as a spacer;[277,278] however, it should be noted that here sample loadability could be an issue, as the capillary volume will be filled for a larger part and resolution might be lost.

STUDY DESIGN

Next to sample collection and handling, the overall study design is a crucial factor in biomarker discovery. Dependent on the research question, the study should be differently designed and different factors need to be taken into account. During this study the aim was to distinguish BPH from PCa patients but, more importantly, to differentiate the different PCa patient groups, indolent from aggressive, specifically in the grey zone of the PSA test (PSA concentration of 3 - 10 ng/mL). For this study a proof-of-principle of the developed PSA Glycomics Assay was shown, however, to validate the outcome, a proper cohort design is needed, ensuring that the patient groups are well matched. Currently, the histology of the biopsy defines the disease status of the patient, however, as stated earlier, tumors can then be missed. Therefore, during the evaluation of a new biomarker, the biomarker should not be compared to the histology of a fresh biopsy but rather on longitudinal information of a patient, preferably data of several years that provides the information if the patient indeed did not harbour PCa at that time or that the diagnosis changed over time. Moreover, the mean average of the total PSA (tPSA) concentration should be similar between the different patient groups and, as age has shown to influence the glycosylation profile in total plasma *N*-glycome,[138,279] it is also important to ensure that the patient groups are age-matched to prevent any age bias. In addition, proper controls need to be taken along, which would preferably be from healthy donors. Unfortunately, as these healthy donors should also be age-matched, they will be mostly above age 60, and there is a good chance that the healthy donors might have a prostate related disease that could bias the outcome of the study. In this case, it is suggested not to have healthy donors as a control sample, but rather to study the glycosylation profiles of the different patient groups and potential differences that are disease specific. Another crucial aspect, which was shown in this thesis, is the biofluid that is selected. While most studies revealed specific glycosylation features for PCa, this was mainly performed on the total *N*-glycome or PSA derived from serum. The study in this thesis provided in-depth information regarding the PSA glycosylation from patients' urine. Even though the obtained results were negative for the proposed hypothesis, it is still suggested to validate the results by repeating the same assay on a different cohort. Preferably a cohort from which the samples were retrieved in a different research group. If similar results can

be obtained, there is a higher probability that the glycosylation of urinary PSA indeed does not contain disease related features. The same accounts for an assay that aims for an altered glycosylation profile of PSA derived from serum between different patient groups. If certain results are found, they should always be validated with another sample set.



CONCLUDING REMARKS

While several studies have already shown the potential of using the glycosylation of PSA as a potential biomarker for PCa, there is still a need for a deeper understanding of this glycoprotein and how the protein and its glycosylation is correlated to PCa. However, to further examine this glycoprotein, technological innovations were needed to enable better sensitivity and repeatability. During these developments questions were raised such as: *What is the mechanism behind the isomer separation of the sialylated glycopeptides (**Chapter 2**)? What is the mechanism behind the improved sensitivity when a dopant enriched nitrogen gas is used (**Chapter 3**)? Does the signal enhancement with dopant enriched nitrogen gas apply to other molecule classes as well (**Chapter 3**)? Can the separation be even further pushed for isomeric separation with CE-ESI-MS (**Chapter 4**)?* Furthermore, the improved CE-ESI-MS platform led to the development of a PSA Glycomics Assay allowing an in-depth study on the urinary PSA ***N***-glycome of patients with elevated PSA levels (> 3 ng/mL) in their circulation, which lead to another question: *Does the in-depth information on PSA glycosylation obtained with the improved CE-ESI-MS platform lead to new biomarkers for PCa (**Chapter 5**)?* As the preliminary results of the **PSA Glycomics Assay** did not show any PCa biomarker candidates in urine, it is suggested that the observed PSA glycosylation profile can be seen as an unbiased glycosylation PSA profile (unaffected by PCa). This observation lead to the following questions: *does the glycosylation of PSA in serum harbor potential biomarkers for PCa? If so, can the unbiased PSA glycosylation profile of urine be used as a personal baseline for each patient?* While, this is just a small selection of all the questions that arose during this project, it provides an indication of how complex research can be and one should be aware that obtaining a single clear answer is most often not the case. Moreover, this thesis provided a deeper insight into the urinary PSA *N*-glycome, which is just the beginning of a longer journey to study and understand how this information could be of potential use to improve the early differentiation of aggressive *versus* indolent PCa. However, it is of utmost importance that a single stakeholder does not get blindsided by his/her own interest and that the overall goal should be clear and valuable for all stakeholders. Therefore, it is crucial in the early stages that the different stakeholders share their knowledge, interests and desires with each other to ensure that indeed the right questions are being asked, crucial factors are being tackled and the efforts will lead in the end to:

Improved cancer patient management