The overall aim of the present proposal was to develop advanced nanoUHPLC-MS(MS) techniques to reliably decipher glycomics and proteomics molecular signatures from prostate cancer tissue microarrays (TMAs) in order to discover key changes occurring during cancer progression in a high throughput manner. This was successfully achieved and is supported by the 8 publications accepted and a further one that is under preparation.

The objective of the **first year** was to develop an HPLC-MS method sufficiently sensitive to analyze limited amounts of glycosaminoglycan (GAG) samples, such as disaccharides released from formalin-fixed, paraffin-embedded tissue biopsies using on-surface enzymatic digestion. The first goal was to develop and optimize the chromatographic separation conditions and to prove the applicability of the developed methods for TMA samples. Based on the literature, a hydrophilic interaction (HILIC) and weak anion exchange (WAX) mixed-mode resin was selected as packing material. The packing material was obtained by opening a commercially available analytical-size column, and then it was repacked into 100 µm i.d. capillaries using a pressure injection cell. The use of capillary columns allows on-line nanoUHPLC-MS coupling thus providing the necessary sensitivity for the investigation of samples consisting of a few micrograms of tissues. The in-house packed capillary column was mounted on a Waters® nanoAcquity UPLC system coupled to a high resolution Waters® Q-Tof Premier™ Mass Spectrometer via a nanoelectrospray ionization source. Method development was performed using commercially available Δ4,5-unsaturated heparan sulfate (HS) disaccharide standards. An isocratic elution method was developed[1], and partially validated, and low LOQ values and good repeatability was observed. The next phase of method development was establishing whether the method would be suitable for analyzing small amounts of biological material. The main challenge of using TMA samples is the limited sample amount (1.5 mm diameter, 5 µm thin cores). TMA cores were digested using heparinase enzymes yielding Δ4,5-unsaturated HS disaccharides. The digestion products were pipetted off the cores, and analyzed by the nanoHPLC-MS method described[1]. TMA cores corresponding to normal, grade 1 and grade 2-3 cancer were chosen in triplicates to assess the repeatability of the developed method. The four most abundant disaccharides in these samples were observed. The most abundant component was the non-sulfated disaccharide (20-80 fmol range in the various samples). The two monosulfated components were present in comparable amounts; 10-40 fmol in the case of *N*-sulfated and 5-20 fmol in the case of *O*-sulfated. Among the higher sulfated components, only the disaccharide carrying both *N*- and *O*-sulfation was observed in each case, but only in low amounts (less than 10 fmol). It was therefore shown that sulfated HS disaccharides can be studied by HPLC-MS using the developed isocratic nano-chromatography conditions[1].

The objective of the **second year** was to characterize the glycosylation pattern of glycoproteins present in prostate cancer TMAs following on-surface tryptic digestion. Glycosylation patterns often change during pathophysiological conditions which is also underlined by the fact that most tumor biomarkers are glycoproteins. Several modifications to existing methods were successfully implemented in order to improve their analytical performance. For this, it was necessary for us to update and slightly modify the Workplan:

1) First, the tissue surface digestion method was thoroughly optimized and a pilot study on prostate cancer TMAs was performed[2]. The method was based on tissue surface tryptic digestion and nano-HPLC-MS/MS measurements, with the application of which we were capable of identifying and quantifying over 500 proteins from 1.5 mm diameter TMA cores. This sensitivity was essential for future glycoprotein analyses. Healthy and cancerous prostate

tissues samples were compared, and tissues with various grades and stages of cancer. Tissue proteomics could distinguish healthy and cancerous samples, furthermore the results correlated well with cancer grade, but not noticeably with cancer stage[2]. These results highlighted the feasibility of the developed method for future large-scale tissue proteomics and *N*-glycosylation studies using commercially available TMAs.

2) For sensitive detection of minor glycoforms in tissues, especially when using very small sample amounts like those available in TMAs, it is essential to include a glycopeptide enrichment step in the analytical workflow. For this purpose, an acetone-based precipitation method was used as this can be easily integrated in the tissue surface digestion workflow. Following initial trials, we have thoroughly optimized the method to achieve both high yield and high enrichment for glycopeptides[3]. Results were tested both on a model glycoprotein, alpha-1-acid glycoprotein (AGP) and a complex HeLa cell lysate. Glycopeptides and peptides were identified by tandem mass spectrometry in the nano-HPLC-MS/MS workflow using the Byonic software. Quantitation of the individual glycoforms was performed based on MS1 measurements, using our in-house developed GlycoPattern software.
In contrast to that stated in the Workplan, AGP standard was used to optimize the acetone precipitation protocol and not the tissue surface digestion itself. The reason was that without an enrichment step, the sensitivity is not sufficient to reliably detect glycopeptides from the tissue surface. That makes the enrichment of glycopeptides necessary, but it is the step most likely to introduce bias in the workflow. Therefore, reproducibility of the enrichment step and the fact that site specific *N*-glycosylation profile was not altered was tested using AGP. Before analyzing the *N*-glycosylation profiles of glycoproteins extracted from the tissue surface we decided to test the enrichment step and glycopeptide analysis workflow on a complex HeLa cell lysate[3]. The tryptic digest of HeLa S3 cells is commercially available and is the most widely characterized mass spectrometry standard in proteomics, however there is only limited information available about the protein *N*-glycosylation profile of the HeLa cervical adenocarcinoma cell line. The reason for analyzing HeLa cells before the TMA samples was their commercial availability and also for initial testing it was much more feasible than analyzing glycopeptides extracted from TMAs where sample preparation is tedious.

3) Another important aspect was developing a nanoHPLC-MS(MS) method suitable for discriminating core and antenna fucosylation[4]. Glycoprotein fucosylation is often used as a biomarker for various diseases, and seems to be extremely important in cancer as well. Fucosylation takes place frequently on the *N*-acetylhexosamine units of *N*-glycans, mostly on the core (connected directly to the peptide chain), occasionally on the antenna. Distinguishing core and antenna fucosylation is often desired but is not trivial. The main bottleneck of the analysis is that fucose is known to migrate along the oligosaccharide chain during mass spectrometry analysis. We developed a straightforward approach to distinguish core and antenna fucosylation in glycopeptides[4]. The method does not require derivatization, and can be easily adapted into a proteomics workflow. The key aspect is to use low energy collision induced dissociation when only single step fragmentation processes occur and interfering processes like fucose migration or consecutive reactions are minimized. Core and antenna fucosylation can be discriminated using various ion abundance ratios[4].

After the second year, the PI was on maternity leave.

The main objective of the **third year** was to apply the tissue surface enzymatic workflow to a larger set of human prostate tissue samples to identify possible new biomarkers of prostate

cancer. To achieve this aim, further improvements to methods were carried out, especially in the case of GAG disaccharide analysis discussed in sections 4) and 5). Glycoproteomics and proteomics were performed on a cohort of TMA samples (n=95), discussed in section 6). Chondroitin-sulfate (CS) and heparan-sulfate (HS) disaccharides were analyzed in a separate cohort of tumorous, and benign prostate hyperplasia (BPH) tissue samples discussed in section 7).

4) The HPLC-MS method developed during the first year for analyzing GAG disaccharides extracted from the tissue surface was further optimized. It was necessary to use a more robust hybrid system combining 250 µm i.d. capillary columns on a nanoHPLC coupled to a mass spectrometer equipped with electrospray ion source. Instead of using isocratic separation, a salt gradient method was developed and partially validated resulting in shorter run times (15 min) and more reliable analysis of small amounts of HS[5] and CS[6] disaccharides. The developed methods enabled analysis of HS and CS disaccharides present in the sub-100 fmol range.

5) GAG disaccharides are labile molecules prone to degradation, therefore stability and recovery of CS-derived disaccharides under commonly used sample preparation conditions were investigated and general sample handling methodologies were proposed[7]. Stability of CS disaccharides with varying degrees of sulfation were studied comparing solvent exchange using vacuum evaporation (SpeedVac) and lyophilization. Use of SpeedVac resulted in higher recoveries and lower standard deviation. Sample loss as much as 20% is expected in the case of aqueous solutions, while from high organic solvent–containing mixtures, losses may increase to 50% due to non-specific interactions occurring between the disaccharides and the tube walls[7]. Next, stability of CS disaccharide mixtures in typical HPLC injection solvents were compared. Disaccharides may decompose during storage in the autosampler, even if it is cooled to 4 °C. In the case of methanol-based solvents, autosampler storage is not recommended for over 12h, since all CS disaccharides decompose or stick to the walls at a high rate. In the case of acetonitrile-based solvents, the decomposition rate of the monosulfated disaccharides (D0a4 and D0a6) presents the worst problem. Storage is best in 50:50 v/v and 75:25 v/v ACN:water, as it results in a bias-free analysis even after 24h[7].

6) Protein and *N*-glycopeptide abundances of 95 prostate TMA samples were determined[8] among which there were 9 grade 1, 16 grade 2, 24 grade 3, and 46 normal tissues. The sample preparation consisted of on-surface tryptic digestion of proteins, extraction followed by C_{18} SPE cleanup, and acetone precipitation for glycopeptide enrichment as developed in the second year of the project. The glycopeptide-enriched pellet and the supernatant fractions containing non-glycosylated peptides were analyzed separately, using a nanoHPLC-MS/MS system. In total, 653 proteins and 145 *N*-glycopeptides were quantified. First, healthy and all the prostate cancer (PCa) tissues were compared and 123 proteins were found to be differentially expressed, 72 upregulated and 51 downregulated[8]. Multiple glycosylation features were altered significantly: the abundance of 7 glycopeptides, and the sialylation, fucosylation, and galactosylation on 9 glycosites. Second, we compared distinct pathological grades and healthy tissue where 75 proteins were differentially expressed, while 4 glycopeptides showed altered abundance and 1 glycosite showed altered fucosylation[8]. The significantly changing proteins were subjected to protein-protein interaction network analysis to reveal information about the biological processes involved. Glycoproteins showing altered *N*-glycosylation were all secreted to either blood or ECM. Our results indicate that alterations between PCa and normal prostate tissue glycosylation occur primarily on the glycosite level, while overall glycosylation may be unaffected. Furthermore, there was no direct correlation between altered glycosylation and differential expression of proteins. Further investigation of the glycosylation, and cancer

specificity of these potential prognostic markers and identification of their exact roles is reasonable and could lead to further advancement in understanding the function of glycosylation in cancer development and PCa prognosis.

7) We investigated the CS and HS content and sulfation pattern of samples originating from patients with benign prostate hyperplasia (BPH) (n=16) and prostate cancer (low (n=20), intermediate (n=16) and high (n=16) risk groups based on Gleason grading. Risk groups were balanced for age and overall survival. The GAG chains of the tissues were degraded into disaccharide building blocks using bacterial lyase enzymes and then purified in a graphite-based pipet tip SPE system. We used the nanoHPLC-MS/(MS) methods described in section 4) for separating and quantifying GAG disaccharides. We have found that neither the overall quantity nor the sulfation motifs of HS chains showed significant changes among the investigated groups. CS chains, however, were altered significantly among BPH and different risk groups of PCa. The overall rate of sulfation was lower in the low-risk PCa than in BPH group, but it increased significantly with the progression of cancer, being especially high in high-risk PCa. Furthermore, the 6S/4S ratio showed a large increase in PCa compared to BPH (1.4-, 1.7-, and 6-fold in low-, intermediate-, and high-risk PCa, respectively), and the changes were statistically significant in all cases except between low- and intermediate-risk PCa. These findings provide a good basis for future potential GAG-based classification of prostate malfunctions. The manuscript describing these findings is under preparation.

References:

1) Turiák L, Tóth G, Ozohanics O, Révész Á, Ács A, Vékey K, Zaia J, Drahos L: Sensitive method for glycosaminoglycan analysis of tissue sections. JOURNAL OF CHROMATOGRAPHY A 1544: pp. 41-48. (2018)
2) Turiák L, Ozohanics O, Tóth G, Ács A, Révész Á, Vékey K, Telekes A, Drahos L: High sensitivity proteomics of prostate cancer tissue microarrays to discriminate between healthy and cancerous tissue. JOURNAL OF PROTEOMICS 197: pp. 82-91. (2019)
3) Turiák L, Sugár S, Ács A, Tóth G, Gömöry Á, Telekes A, Vékey K, Drahos L: Site-specific N-glycosylation of HeLa cell glycoproteins. SCIENTIFIC REPORTS 9(1):14822. (2019)
4) Ács A, Ozohanics O, Vékey K, Drahos L, Turiák L: Distinguishing core and antenna fucosylated glycopeptides based on low energy tandem mass spectra. ANALYTICAL CHEMISTRY 90(21): pp. 12776-12782. (2018)
5) Tóth G, Vékey K, Drahos L, Horváth V, Turiák L: Salt and solvent effects in the microscale chromatographic separation of heparan sulfate disaccharides. JOURNAL OF CHROMATOGRAPHY A 1610: pp. 460548. (2020)
6) Tóth G, Vékey K, Sugár S, Kovalszky I, Drahos L, Turiák L: Salt gradient chromatographic separation of chondroitin sulfate disaccharides. JOURNAL OF CHROMATOGRAPHY A 1619: pp. 460979. (2020)
7) Tóth G, Pál D, Vékey K, Drahos L, Turiák L: Stability and recovery issues concerning chondroitin sulfate disaccharide analysis. ANALYTICAL AND BIOANALYTICAL CHEMISTRY 413: pp. 1779–1785. (2021)
8) Sugár S, Tóth G, Bugyi F, Vékey K, Karászi K, Drahos L, Turiák L: Alterations in protein expression and site-specific N-glycosylation of prostate cancer tissues. SCIENTIFIC REPORTS 11(1): pp. 15886. (2021)