Summary

N- and O-glycosylation can modulate structural and functional properties of glycoproteins and is involved in regulating cellular processes. Consequently, altered glycosylation patterns have been associated with numerous diseases including inflammation and cancer, making glycosylation a very promising tool for clinical diagnostics, prognostics and therapeutics. Protein glycosylation is one of the most challenging analytical targets due to its vast structural heterogeneity and dynamic nature. However, the ability to perform detailed structural characterization of the glycome and glycoproteome is the basis to investigate and subsequently understand the biological functions of protein glycosylation as well as its value for clinical applications. The development of new methods and the improvement of existing ones are necessary for this purpose.

Along that line the first part of this thesis focuses on the investigation of glycopeptide ionization, fragmentation and separation in order to establish workflows for detailed site-specific glycopeptide analysis in combination with software tools for automated compound identification.

In Chapter 3 the ionization efficiency of glycopeptides was investigated using a defined set of synthetic glycopeptides and their corresponding unglycosylated (Asn) and deglycosylated (Asp) peptide standards. Equimolar mixtures of these (glyco)peptides varying only in their glycosylation site position within the peptide were analyzed by micro- and nanoflow ESI-IT-MS, ESI-QTOF-MS, ESI-FT-ICR-MS, MALDI-FT-ICR-MS and MALDI-TOF-MS. In all MS setups the analyzed glycopeptides showed lower signal intensities than their equivalent peptides, evidently resulting from lower ionization efficiencies. Depending on the localization of the glycosylation site relative differences in signal intensities were also observed between unglycosylated and deglycosylated tryptic peptides, in particular when singly or doubly charged ions were formed. These are important findings to set reliable standards for label-free quantitative glycoproteomics.

In Chapter 4 glycopeptides from the same set of standards were used to optimize glycopeptide fragmentation on QTOF mass spectrometers. Parameters were tuned for a combined lower- and higher-energy CID (stepping-energy CID) of glycopeptides to achieve optimal sequence coverage of both the peptide and glycan portion in a single fragmentation spectrum. The method was evaluated by glycopeptide analyses performed on all human immunoglobulins in combination with automated glycopeptide identification.

The knowledge gained from the work described in these two chapters was applied to develop a LC-MS based method for glycopeptide analysis involving C18-RP and PGC-LC within a single experiment. PGC-LC-MS is mainly used for glycan analysis due to its separation power for glycan isomers. A review about the application of PGC-LC-MS for clinical glycomics is given in Chapter 2. Furthermore, in Chapter 5 PGC-LC
was coupled to C18-RP-LC for an integrated C18-PGC-LC-ESI-QTOF-MS/MS approach using stepping-energy CID for in-depth N- and O-glycopeptide analysis. Pronase-treated samples were analyzed, resulting in glycopeptides with a small peptide portion. The combination of the two stationary phases allowed the capturing of more hydrophobic as well as hydrophilic compounds, revealing a more comprehensive glycopeptide characterization.

The second part of this thesis includes the application of these methods for in-depth glycoprotein characterization of IgG3 N- and O-glycosylation, in particular the hitherto undescribed C³-domain N-glycosylation site (Chapter 5). Using C18-PGC-LC-ESI-QTOF-MS/MS Pronase-generated N-glycopeptides of IgG3 C³-domain, C⁴-domain and hinge-region O-glycopeptides were identified in a single experiment without prior enrichment. Remarkably, IgG3 N-glycosylation is domain-specific featuring exclusively nonfucosylated glycans for the C⁴-domain with a high degree of bisection compared to the C³-domain.

C18-PGC-LC-ESI-QTOF-MS/MS was also used for the detailed N- and O-glycosylation characterization of the heavily glycosylated plasma glycoproteins C1-Inhibitor (Chapter 6). In combination with a panel of other MS-based approaches C1-Inhibitor released N-glycans as well as N- and O-glycopeptides were investigated. The total C1-Inhibitor N-glycome was determined using the PNGase F released N-glycans, complemented by a site-specific glycopeptide analysis. All six N-glycosylation sites were successfully identified and for most of them the relative distribution of the corresponding glycoforms could be determined. Further 11 O-glycosylation sites were identified carrying core 1-type O-glycans, next to a heavily O-glycosylated region spanning from Thr82-Ser121 that contained up to 16 occupied O-glycosylation sites.

The established glycoproteomics workflow was further applied for the analysis of a Fc-fusion protein, showing the potential of this method in early biopharmaceutical development (Chapter 7). The Fc-fusion protein contains one N-glycosylation site, which was characterized on a level of Pronase- and trypsin-treated glycopeptides using LC-MS as well as by released N-glycan analysis using MALDI-TOF-MS after linkage-specific sialic acid derivatisation. Furthermore, two O-glycosylation sites were identified and characterized using an ETD and stepping-energy CID approach on Pronase-treated glycopeptides.

Whereas in the previous chapters the focus was the in-depth characterization of purified glycoproteins, more complex clinical samples were the target of the experiments described in Chapter 8. The released N-glycans of human bowel tissue biopsies were characterized along the bowel from the terminal ileum to the rectum of control and ulcerative colitis patients. In total, 143 N-glycan structures were identified using a PGC-LC-ESI-MS/MS glycomics approach. Relative quantitation was also performed on 120 of the released N-glycan alditols. Annotated N-glycan spectra were used to establish a bowel N-glycome fragmentation spectra library, which also built the basis for ongoing large-scale studies. Additionally, we investigated region-specific glycosylation features
along the bowel to set standards for biopsy sampling. A particularly striking difference in the N-glycosylation patterns was found between the terminal ileum and the entire colon, but hardly any region-specific patterns could be identified within the large intestine.

In the last chapter, Chapter 9, current technologies, opinions and future trends regarding glycan and glycopeptide separation, identification and the application of glycopeptide standards are discussed within the context of this thesis.