High-throughput mass spectrometric N-glycomics

Reiding, K.R.

Part IV – Discussion

Chapter 9

General discussion

General discussion

While glycosylation is an important co- and post-translational modification exhibited by more than 50% of proteins[1,2], the analysis thereof is a remarkably complex affair. Underlying challenges include the non-linear arrangement of a varying number of monosaccharides, an absence of fluorescence response of native glycans, as well as heterogeneity in glycoprotein characteristics such as solubility, charge and stability[3].

This thesis mainly focuses on the method development and application within one particular branch of glycosylation analysis methodologies, namely matrix-assisted laser desorption/ionization mass spectrometry (MALDI) mass spectrometry (MS). For analysis of the ions we have made use of either time-of-flight (TOF) or Fourier-transform ion-cyclotron resonance (FTICR). The major advantage of MALDI over other ionization methods like ESI is the speed by which samples can be measured – the acquisition time for an individual sample is typically only a few seconds – making it particularly suitable for the mass spectrometric analysis of large cohorts[4-6]. What remained, and was explored in the work covered in this thesis, was to improve information content and quality, to decrease wet-lab hands-on time, to handle glycomics data integration, to streamline normalization, derived trait calculation and statistical analysis, and to apply the methodologies to challenging and informative clinical cohorts.

9.1 Developments in linkage-specific sialic acid esterification

In this thesis, two overall derivatization strategies are described for released *N*-glycans, namely the esterification of the acidic residues present on sialylated species (**Chapter 3**), and reductive amination of all *N*-glycans at their reducing ends (**Chapter 6**). The first led to the stabilization and linkage-discretization of sialic acids (both of the *N*-acetylneuraminic acid and *N*-glycolylneuraminic acid variety; **Chapter 3** and **Chapter 5**), while the second allowed high-resolution MALDI-FTICR-MS in negative mode (**Chapter 6**).

9.1.1 Application of linkage-specific sialic acid derivatization

The esterification protocol proved beneficial for the analysis of released *N*-glycans from human serum and plasma, and has contributed to a variety of case-control and longitudinal studies including on Down Syndrome[7] and pregnancy[8-10]. Furthermore, the reaction could be applied to the study of antibody glycosylation, examples including the released *N*-glycans of IgG-Fc and IgG-Fab[8], and those from engineered Her2-IgAs[11,12]. Direct comparison studies revealed that MALDI-TOF-MS with relatively simple mixtures of esterified *N*-glycans showed largely similar profiles and performances as liquid chromatography with fluorescence detection (**Chapter 2** and **Chapter 3**)[13-16]. Furthermore, our method comparison on the analysis of total serum *N*-glycosylation (TSNG) (**Chapter 8**) showed a generally higher information content (more individually detected signals) with MALDI-TOF-MS than with HILIC-UHPLC-FLD and xCGE-LIF. When applying the workflow to the analysis of murine

plasma, known to contain various glycan structures different from those found in humans[17,18], we could also establish the similarity of the reaction on α2,3- and α2,6-linked *N*-glycolylneuraminic acids, the esterification of sialic acids linked to *N*-acetylgalactosamines instead of galactoses, as well as the qualitative preservation of sialic acid *O*-acetylation due to the mild reaction conditions (**Chapter 5**)[19,20].

A downside of the ethyl esterification turned out to be the long-term stability of the esters, particularly the lactone[21], and while full reaction selectivity of the carboxylic acid modification was achieved for sialic acids, experiments on IgG glycopeptides revealed a variety of reaction products on the peptide portion[22,23]. To suit specific purposes, alternative reaction variants were designed, such as dimethylamidation / lactonization for IgG glycopeptides (showing full reaction specificity on the tryptic IgG peptide portion)[22,24], and dimethylamidation / amidation for increased long-term stability required for tissue imaging (**Table 1**)[25].

Table 1. *Sialic acid masses before and after linkage-specific derivatization.*

Modification	Nucleophile	Solvent	Δmass	Neu5Ac Δmass	Neu5Gc Δmass	Refs
α2,3-linkage-specific						
Lactonization	-	Methanol; Ethanol; DMSO	-18.0106	273.0849	289.0798	16,26
Amidation	Ammonium hydroxide	DMSO	-0.9840	290.1114	306.1063	25
Methylamidation	Methylamine	DMSO	13.0316	304.1271	320.1220	27,28
α2,6-linkage-specific						
Methyl esterification	Methanol	Methanol	14.0157	305.1111	321.1060	16,26
Ethyl esterification	Ethanol	Ethanol	28.0313	319.1267	335.1216	16
Dimethylamidation	Dimethylamine	DMSO	27.0473	318.1427	334.1376	22,25
Non-derivatized						
Native	-	-	-	291.0954	307.0903	-
Sodiated [M-H+Na]	-	-	21.9819	313.0774	329.0723	-
Potassiated [M-H+K]	-	-	37.9559	329.0513	345.0462	-

The ethyl esterification protocol, encompassing the chemical derivatization of glycans, HILIC enrichment and MALDI-TOF-MS measurement, proved automatable, but required modifications in the reaction temperature, the HILIC stationary phase, and the matrix solvents (**Chapter 4**)[29]. The resulting shift of the analytical bottleneck to data analysis was addressed by the development of software packages for integration and quality control,

such as the recently-published MassyTools[30]. Notably, the data analysis tools could make use of the enhanced resolution of the now-attained reflectron-mode MALDI-TOF-MS measurements. This includes quality control like high-precision ppm error determination on the now well-isolated monoisotopic signals, and comparison of theoretical with observed isotopic envelopes. The resulting high-throughput *N*-glycomics workflows were applied to several of the larger mass spectrometric total serum *N*-glycomics studies to date (**Chapter 6** and **Chapter 7**)[31].

9.2 Persistent challenges in mass spectrometric *N*-glycomics

While this thesis describes advances in MALDI-MS *N*-glycomics throughput and structural information (sialic acid linkage), the methodology is at current not capable of unravelling the full complexity of glycosylation in biological systems. Taking human glycosylation as an example, several layered challenges still need to be addressed. These include a larger control of technical variability, an increase in structural information, a traceability of glycans to specific glycosylation sites, as well as a coupling of glycomics with the analysis of other post-translational modifications (ideally in a protein-specific manner).

9.2.1 Controlling technical variability

So far, identification of glycomics biomarkers has been made difficult by relatively low correlation coefficients, indicating a lot of uncontrolled variability in the data. While part of this may arise from biological sources (which need to be understood), a large contribution may come from measurement errors as well. Most chapters in this thesis describe coefficients of variation based on positive controls in the experimental run (**Chapter 3-8**)[16,19,29,31], typical values ranging from 1% to 6% for the highest signal but increasing with decreasing signal intensity. While quite decent from an analytical perspective, this nonetheless means that additional effort may be useful to reduce this component of the cumulative error.

One complicating factor in *N*-glycan analysis is the variable abundance of the reducing-end glycosylamine which is present for glycans upon PNGase F release and converts over time to a regular aldose in aqueous conditions[32]. Performing the ethyl esterification protocol in the presence of these glycosylamines can lead to the unwanted reducing-end labeling with EDC or a fragment thereof, which in turn diminishes the signal intensity of the desired product. Fortunately, variability in PNGase F digestion and glycosylamine decay may be reduced, for example by increasing protocol incubation times, as reactions will principally be most robust when performed to completion. However, elongating incubation steps also leads to a conflict with method throughput. Recent reports describe attractive reducing-end labeling methods that make use of rapid PNGase F digestion and subsequent glycosylamine labeling, with complete deglycosylation of samples (IgG) reported within 10 minutes[32,33]. However, uniform glycosylamine production will presumably be hard to

achieve for all glycan species in a complex mixture – the most stable end-product being the aldose – making the strategy particularly risky when following up with ethyl esterification.

Another source of analytical variability lies in the sample purification step. Sample purity is often a requirement for mass spectrometric glycomics, as glycans (and glycopeptides) show a large ionization disadvantage over non-glycosylated peptides, and may be ion-suppressed in case of mixtures[34]. In addition, many release and derivatization protocols introduce high concentrations of reagents and detergents, preventing MS analysis of the crude reaction mixture[16,35,36]. One way to ensure the necessary sample purity is by performing a step of solid-phase extraction (SPE). In glycan analysis this typically involves making use of stationary phases (*e.g.*, C18-functionalized beads, PGC, a sugar polymer), and a mobile phase comprising water, an organic solvent, or a mixture thereof[31,37,38]. However, using SPE comes with several challenges as well. It is very difficult to keep the mobile phase or elution buffer concentrations constant during a multiple-sample SPE-session, as selective evaporation of volatile organics will inevitably occur. Unfortunately, even small changes in the organic solvent concentration of a mobile phase or elution buffer will influence glycan recovery in a species-dependent manner. Also challenging for this reason is elution, as residual washing buffer present on a stationary phase will be variable and cause differences in analyte recovery from one sample to the next.

With manual preparation it is possible to make small adjustments to a protocol to compensate, *e.g.*, keeping solvents at cold temperature and in closed bottles to limit selective evaporation or applying a slightly longer vacuum step when residual solvent is visibly present, but for an automated protocol (**Chapter 4**) every step needs to be predefined (and is immutable). As such, to allow sufficient repeatability for an automated variant of the ethyl esterification protocol (**Chapter 4**), the established workflow (**Chapter 3**) had to be adjusted in several ways. Evaporation was limited by covering the fluid containers with plastic foil (to be pierced at the first pipetting step) and by placing two covers on each reaction plate when incubating. Vacuum washes at concentration-sensitive steps were elongated to limit solvent residuals, and elution after SPE was performed with an automated centrifuge. One particular head-ache of protocol automation, namely the formation of bubbles in mixing steps, was prevented by increasing organic solvent concentrations, most notably of the MALDI matrix solution.

A profoundly different method of clean-up may lie in the covalent immobilization of glycans or glycoproteins to a stationary phase, of which applications have already been presented[39-41]. An advantage of this is the less strict requirement of washing solution organics concentrations, and the protocols can be directly combined with reducing-end labeling of the glycans[41]. Having said this, covalent methods of purification have not yet been reported to have lower variability than SPE methods of purification, indicating that further optimization might still be necessary[39-41].

An interesting observation throughout our MALDI-MS studies is that derived glycosylation traits, biology-inspired mathematical expressions encompassing chemically similar glycan species, show a lower technical variation than the single glycan traits (**Chapter 4** and **Chapter 8**)[29]. Next to grouping by chemical similarity, derived traits cluster *N*-glycan species that are found in a similar *m/z* range, *e.g.*, triantennary species (± *m/z* 2700 – 3200 when esterified $[M + Na]^+$) and tetraantennary species (± *m/z* 3200 – 3800). In addition, the derived traits cluster species with similar intensity, for example the di-, tri, and tetraantennary species which differ in intensity by approximately an order of magnitude [**Chapters 3**, **4**, **6**, and **8**][16,29,31]. Part of the uncontrolled variation may thus be caused by sample preparation aspects that differentially affect the chemically divert (and massed) glycan structures, or by the limited dynamic range of the MS instrumentation (*i.e.*, the range in which analyte concentrations show a linear relationship with signal intensity)[6]. For the derived traits, compared to the native glycans, the influence of both sources of variation would then be limited.

One interesting solution to limit variation on the experimental side of MALDI-MS could be the inclusion of internal standards to a sample, for instance by heavy-isotope reducing-end labeling[42-44], permethylation[15,45], or sialic acid derivatization[46]. In addition to decreasing the dependence on mathematical adjustments such as the commonly employed total area normalization and derived trait calculation (**Chapter 5**), this internal standardization would provide a (semi-)absolute measure of analyte quantity.

9.2.2 Stereoisomerism

As exemplified by **Chapter 8**, MALDI-MS principally does not discriminate between glycan compositional and constitutional isomers. The differences in monosaccharide enantiomers (D and L), diastereomers (mannose and galactose), anomers (α and β), and linkages (usually 1,3, 1,4 or 1,6, with 2,8 and 2,9 possible for sialic acids) therefore stay invisible. For screening experiments, MALDI-MS is limited to providing the monosaccharide composition in a number of hexoses, *N*-acetylhexosamines, deoxyhexoses, and sialic acids (*e.g.*, H5N4F1S2). Nonetheless, several chapters have made assumptions on *N*-glycan structural features (**Chapters 5**, **6**, and **7**), primarily based on literature knowledge gathered with, *e.g.*, nuclear magnetic resonance spectroscopy, lectin interaction assays, and a variety of glycoanalytical methods combined with exoglycosidase digestion[47-52]. For example, throughout this thesis the composition H5N4 has been interpreted as a structure with all HexNAcs being GlcNAcs and the five hexoses comprising three mannoses and two galactoses, *i.e.*, a diantennary glycan with two galactoses. However, for this composition a biologically valid alternative is a hybrid structure without galactoses and two antennary GlcNAcs. In addition, one of these HexNAcs might be a bisected GlcNAc, or instead be a GalNAc in a LacdiNAc motif (GalNAcβ1,4-GlcNAc). Because of this, a derived trait describing the “galactosylation of diantennary species” may be confounded by bisection, hybrid-type

glycosylation, and any other type of structure with the same monosaccharide composition. Part of the scientific process is to build on previously established knowledge, and structural presumptions are necessary to efficiently produce the link between profiling experiments and functional studies. However, for any given experiment, particularly when performed on samples for which limited prior knowledge is available, these presumptions may turn out to be false.

Mass spectrometry does provide several opportunities to characterize glycans on a structural level. Reports have been made on diagnostic ions found in negative mode MS/MS (*e.g.*, the D-ion)[53], MS/MS of sodium adducts of permethylated glycans[54,55], and differentiation between GlcNAc and GalNAc in high-energy fragmentation[56]. In addition, MS can be hyphenated to various forms of chromatography, CE, and ion mobility for an extra experimental dimension that can benefit structural interpretation[57-59]. However, while hyphenation and orthogonal analysis can usually pinpoint the major structures for a given signal, in complex mixtures the ratios may be hard to relatively quantify, and for large sample sets the workload may be too much to handle. Typically, thorough analysis is performed on a limited set of samples from a cohort (or a pool thereof), which then provides a possibility-space for annotation.

9.2.3 Site-specific glycosylation analysis

The associations between total plasma (or serum) *N*-glycosylation and other clinical variables reported throughout this thesis (**Chapters 6** and **7**) can often be explained by a change in plasma glycoprotein levels rather than glycosylation changes within a specific glycoprotein, but very little objective discrimination can be made between the two[31]. The analysis of glycopeptides rather than released glycans would provide protein- and site-specificity, but while mass spectrometry is a technique that is principally capable of doing this, there is still a lot of ground to cover. Current glycopeptidomics approaches can be grouped into two overall strategies, namely a glycomics-based variant that attempts optimal recognition of distinct glycoforms for a small set of peptides, or a proteomics-based variant that attempts optimal identification of the peptides but usually provides low information on the glycan portion.

For the first approach, glycomics-based, a protein of interest is typically isolated from a complex mixture, digested with trypsin, LysC or another protease to yield distinct peptides, some of which are glycopeptides[60-63]. These glycopeptides can then be analyzed by MALDI-MS, LC-MS or other methodologies. This approach has for example been extensively used for the analysis of tryptic IgG-Fc glycopeptides, making distinction between different subclasses of the antibody (IgG1-4)[60,64]. Tandem MS has shown effective in obtaining additional structural information in these situations. Ion trap instruments may, for example, make use of collision-induced dissociation (CID) and electron-capture dissociation (ECD) to fragment the peptide portion of glycopeptides, or of electron-transfer dissociation (ETD) to

preferentially fragment the glycan portion[65,66]. Comprehensive glycopeptide information may be obtained, among other methods, by a combination of the above[67], by the application of multiple collisional energy levels[63], by the higher-energy collisional dissociation (HCD) methodology that is available for C-trap-equipped Orbitrap instruments[68], or by negative mode MSn approaches[69]. Next to this, a variety of derivatization strategies have been proposed to enhance separation of the glycopeptides, for example by derivatization to allow linkage-specific sialic acid analysis, to enhance ionization efficiency, or to allow isobaric labeling[22,70,71].

The second overall strategy, proteomics-based, attempts to broaden the scope of peptides that can be detected, but for the glycan part usually ends at binary recognition of site-occupancy[72-74]. Reported approaches include proteolytic digestion of the complex mixtures of proteins (*e.g.*, blood plasma, spinal fluid) and subsequent endoglycosidase treatment (PNGase F, PNGase A, Endo H)[72-74]. Readout of the location of glycosylation is provided by the asparagine deamidation as produced by PNGase F or PNGase A after glycolysis, or the detectable presence of a single (fucosylated) *N*-acetylglucosamine as provided by endoglycosidases H or F[72-74]. As is clear, this strategy may lead to the identification of glycosylation sites, and in some instances may be indicative of the core-fucosylation of these sites, but no further information on microheterogeneity is provided[74].

9.3 Developments human plasma *N*-glycosylation analysis

However limited our understanding of human *N*-glycosylation may seem to be, we have also made several advances in the processing and interpretation of complex mixtures like the *N*-glycosylation of blood plasma and serum.

9.3.1 Mathematical developments

For the serum/plasma *N*-glycosylation we have analyzed throughout the thesis (**Chapters 3-8**), we always represent each signal as a fraction of the total analyte area (total area normalization). This form of normalization, also used for UHPLC-FLD and CGE-LIF glycomics, is highly beneficial for the repeatability and data quality of a run[75,76]. However, the downside of such an approach is the introduction of potentially spurious correlation, as the relative increase in one peak will mathematically lead to a decrease in all the others. This conversion works remarkably well for samples with a defined number of glycosylation sites, *e.g.*, IgG-Fc with one *N*-glycan site per C_H2 domain at Asn297[77,78], as the negative correlation between glycoforms will approach reality; for a single glycosylation site the higher abundance of one species necessitates the lower abundance of the others. However, in case of multiple proteins, such as found in plasma, the number of glycosylation sites will differ by the abundance of those glycoproteins. This means that for total plasma *N*-glycome (TPNG) analysis changes in glycosylation values are representative of both changes in protein glycosylation and glycoprotein abundance.

With a reasonable understanding of glycan processing[3,79,80], it is possible to create mathematical constructs to express this in TPNG data, as explored in **Chapter 4** and used within each subsequent chapter. For example, one can relate a nongalactosylated glycan species to a galactosylated species, and thereby obtain a ratio of galactosylation for that species. However, the situation is complicated by the fact that glycans are not independent, but related to many other glycans. For instance, while a nongalactosylated glycan may be processed into a monogalactosylated variant, the latter might be further processed into a digalactosylated galactose thereby hiding the change in the monogalactosylated one[3,74,75]. This means that single enzymatic events may not always be reflected in the direct ratios of product and educt, and the need arises to see glycan processing in a wider framework. An example of a derived trait taking this into account is the calculation of the "degree of galactosylation within all diantennary glycans", yielding a higher number for each additional galactose.

Furthermore, each monosaccharide addition (or removal, but let us regard addition) is dependent on the presence of an appropriate substrate, *e.g.*, galactosylation on the availability of an appropriately linked *N*-acetylglucosamine, and sialylation on the presence of a galactose[3,79,80]. This dependency may also influence biological insight into disease mechanisms. For example, sialylation associates with RA disease activity while sialylation-per-galactose does not, implying that galactosylation and not sialylation might be the important factor to consider in terms of dysregulation[81]. For the derived traits it then is logical to consider these dependencies as well.

In addition, with an estimate of the contribution of plasma proteins to the observed total plasma *N*-glycome[47,48], theoretical protein groupings of glycosylation can be designed. The most prominent example of this is the glycosylation of the Fc portion of IgG, which, to our knowledge, is by far the main contributor to the nonsialylated complex glycans[48]. Looking within plasma at a derived trait describing the galactosylation of nonsialylated glycans (supposed to be mainly IgG-Fc), we see a remarkable congruency with the effects described for IgG itself(**Chapter 6** and **7**)[81,82].

The concept of derived traits has been prior explored in literature, mostly for the HPLC analysis of IgG[83,84]. This protein contains an envelope of glycosylation consisting of glycan species that are largely well-separated by chromatography, which allows for the construction of reasonably well-defined traits[83,84]. However, as demonstrated in **Chapter 8** of this thesis, the chromatographic separation of the *N*-glycosylation from plasma is more structurally ambiguous, particularly in the high antennarity regions of a profile, and poses some challenges for the construction of derived traits. MALDI-MS, while lacking in structural information (*e.g.*, core- *vs.* antennary-fucosylation, α1,3- *vs.* α1,6-branch galactosylation), does show for plasma *N*-glycans an excellent separation on a compositional level and can therefore benefit from derived traits based on these compositions. **Chapters 5** through **8**

describe a large number of compositionally derived traits for serum/plasma *N*-glycosylation analysis, in some cases outnumbering the single observed species. Favorably, these derived traits performed very well for the analysis of plasma *N*-glycosylation, displaying lower methodological variability, yielding stronger statistical associations, and overall providing information that is better clinically interpretable than single *N*-glycan compositions.

9.3.2 Serum and plasma protein grouping

As it is, several broad protein categories were suggested throughout this thesis to influence the serum/plasma *N*-glycome (**Chapter 6** and **7**)[31]. Most clearly we can discriminate a set of glycans also observed on the Fc portion of IgG, secondly we can detect the glycans commonly found on other plasma-cell-produced immunoglobulins together with IgG-Fab, thirdly we can identify the glycosylation also detected on the plasma-abundant acute-phase proteins, and finally we seem to isolate the glycosylation of lipoproteins reflecting lipid transport[31,47,48].

When compared to the overall *N*-glycan structures found in human plasma, the Fc-portion of IgG shows complex-type species with a rather low state of post-translational processing. Whereas most complex *N*-glycans have a near-complete (sub-)terminal galactosylation and sialylation[31,47,48], the *N*-glycans on IgG-Fc typically shows incomplete galactosylation and around 20% of structures carrying sialic acids[8,77]. Plasma cells, the cells that produce a major part of the IgG found in plasma[77,85], are fully capable of high degrees of galactosylation and sialylation. This is apparent from the larger diantennary *N*-glycans observable on IgM, IgA and other plasma-cell derived immunoglobulins[47,48,86,87], from the glycans expressed on the Fab portion of IgG[8], and also from reports on *in vitro* stimulation of B cells to achieve altered Fc-glycosylation[88]. In all, the evidence indicates a protein- and site-specific regulation of the glycosylation of IgG-Fc, which is of biological relevance due to the large influence of the carbohydrate on Fc-gamma receptor (FcγR) binding[77,89]. Molecular modelling of IgG reveals that the Fc-glycans of IgG are inwards oriented, meaning that the low complexity may arise from steric hindrance, *i.e.*, a decreased accessibility of the glycan for glycosyltransferases[90,91]. Another interesting hypothesis is that the glycan may interact with the peptide backbone at specific stages of processing[92]. For example, as soon as the α1,6-branch *N*-acetylglucosamine is galactosylated this residue may form hydrogen bonds with a nearby lysine (Lys246) and threonine (Thr260), thereby limiting further sialylation[92-94]. By whichever means the glycosylation of IgG-Fc is generated, it is sufficiently different from the glycosylation of other abundant plasma proteins that its influence can be seen when looking at glycosylation subgroups in the total plasma/serum *N*-glycome[47,48]. For example, the decrease of IgG-Fc galactosylation with age is a repeated observation with a variety of analytical platforms[83,95,96]. Using the IgG-Fc-specific derived trait in our healthy human plasma study (**Chapter 6**), we could detect the strong association with age as well.

The other main contributors to the plasma/serum *N*-glycome are the acute-phase proteins, coagulation factors (in which lies the difference between plasma and serum), and lipid transport proteins[47,48,97]. Fibrinogen, transferrin, haptoglobin, alpha-1-antitrypsin and alpha-1-acid glycoprotein, all involved in coagulation and the acute-phase response, display typically di- to tetraantennary glycan structures with near-complete galactosylation and sialylation (both α2,3- and α2,6-linked), and show a preference of antennary- over core-fucosylation (in case of α2,3-linked sialylation, thereby forming the glycan epitopes sialyl-Lewis X or A)[16,48,49,98-101]. A common observation in our plasma/serum *N*-glycome studies (**Chapters 6**, **7** and **8**) is that injury and inflammation lead to an altogether upregulation of tri- and tetraantennary glycan species, a phenomenon that would be most easily explained by an upregulation of the abundance of the aforementioned acute-phase proteins, but regulation of glycan microheterogeneity will likely also play a role.

The last plasma protein grouping observed in our studies involved lipid transport (**Chapter 6**)[47,48]. Of interest is that the larger high-mannose glycan structures (*e.g.*, Man8 and Man9) are not commonly found on glycoproteins that contribute to the total plasma *N*-glycome[47,48]. In fact, the only protein we have identified to have sufficient influence on the Man9 content in the TPNG is apolipoprotein B-100, a structural protein in the formation of (V)LDL particles[102-104]. In a fascinating parallel, our study regarding metabolic markers in healthy individuals (**Chapter 6**) shows one of the strongest associations between glycosylation and lipid metabolism to be with the size of high-mannose glycans, notably by the upregulation of Man8 and Man9. The simplest biological explanation of this association would be the increased abundance of LDL particles and accompanying high-mannose-glycosylated apolipoprotein B-100, but also here specific regulation may play a role.

In all, it has to be noted that many of the prior statements have been logically induced from joining the known protein concentrations in human fluids with the overall understanding of the glycosylation thereof[47,48], but this may of course be untrue in any particular case. An important step for future research would be the actual quantification of the proteins from plasma, combined with the glycosylation thereof. This may also provide information on whether our repeated observations come from the change in glycoprotein quantity or from an altered expression of protein glycoforms. Furthermore, not all high-abundant plasma protein could be included in the theoretical calculation. For instance, there are several acute phase proteins that are, by the discrepancy between their conceptual protein mass and actual observed mass, likely carriers of high degrees of tri- and tetraantennary types of glycosylation. Examples of this are histidine-rich glycoprotein and kininogen-1[105,106], but so far these have not been characterized with respect to glycan macro- and microheterogeneity[48].

9.3.3 Cellular origin of plasma *N*-glycosylation

Altogether, it appears highly likely that the total plasma *N*-glycome represents the glycosylation characteristics of a limited set of cell populations, the main contributors being plasma cells and hepatocytes. These populations appear to have discrete glycosylation-type outputs on their secreted glycoproteins[48]. For instance, plasma cells are rather unique in their abundance of core-fucosylated *N*-glycans, their processing of glycans to at best diantennary complex-type, the near-exclusivity of α2,6-linked sialylation, and the occurrence of bisection[48]. Hepatocytes, on the other hand, readily show the capability to generate tri- and tetraantennary *N*-glycans, LacNAc repeats, antennary fucosylation, both α2,6- and α2,3-linked sialylation, and so far no bisected structures have been identified. For cell-line engineering it was previously shown that the cell type has a large influence on the glycosylation type that can be expected. Examples of this are CHO cells and HEK cells, each providing their unique glycosylation characteristics[107,108]. Protein-specific attributes such as steric hindrance and glycan-protein interaction can then fine-tune the glycosylation per site, but the overall regime appears defined by the cell type.

Nonetheless, plasma will also contain glycoproteins originating from other cell populations, such as kidney cells and thyroid cells[109,110]. Of most clinical interest would be the plasma glycoproteins that originate from aberrant cell populations, such as tumor tissue. With sufficient analytical sensitivity, it may be possible to see the minor amounts of uncommon glycan structures that originate from these pathological tissues, potentially providing the simple biomarker with high sensitivity and specificity that is coveted by clinical practice. The most promising biomarkers are commonly of low abundance[111], signifying the need for analytical tools of high sensitivity.

9.3.4 Clinical associations within the plasma/serum *N*-glycome

Chapters 6 and **7** have explored the relationship between plasma as well as serum *N*-glycosylation and physiological and clinical parameters like age, sex, lipid levels, markers of inflammation, and a variety of others. These studies are not the first ones to associate glycosylation with clinical variables, but are exceptional by the large-scale use of mass spectrometry, which provides information orthogonal to that of other glycoanalytical methods (**Chapter 8**).

Several total plasma *N*-glycosylation traits show to differ between the sexes. One prominent example of this is tri- and tetraantennary fucosylation (presumable antennary-bound in the form of sialyl-Lewis X), which show on average to be higher in males than in females[31,112,113]. While it is difficult to attribute these glycosylation differences to one cause or another, one attractive direction of future research may be hormonal regulation, *e.g.*, by estrogen as indicated in the literature[114,115]. Association with hormone levels may also be the driving force behind several of the pregnancy-associated glycosylation changes we observed in the total serum *N*-glycome throughout pregnancy (**Chapter 7**).

Other clear associations seen within the total plasma *N*-glycome appear to largely stem from various aspects of inflammation. **Chapter 6** confirmed with mass spectrometry the strong negative association between calendar age and galactosylation of IgG-Fc-type *N*-glycans[31,83], a widely-reported phenomenon thought to be related to the build-up of chronic inflammation with age[95]. Furthermore, as prior mentioned, a large part of the tri- and tetraantennary structures within plasma stem from acute-phase proteins that differ in level upon inflammation[116], and may alter their glycosylation to affect their function and recognition[116,117]. However, when assessing the biomarker potential of glycosylation for the diagnosis of specific diseases, it has to be kept in mind that these markers of general inflammation will likely be picked up as well and will have limited specificity in clinical screening.

Furthermore, lipid metabolism, marked in its simplest form by BMI, but also by measures of, *e.g.*, cholesterol, leptin, adiponectin, and triglycerides, show in part the same glycosylation associations as inflammation, but show in addition also several unique effects (**Chapter 6**). One particular glycosylation characteristic, highlighted in the metabolic study of **Chapter 6**, is the glycan composition H4N4S1 which provides discrimination between healthy and unhealthy metabolic profiles, and is an interesting target for future research.

9.4 Future prospects of plasma *N*-glycosylation analysis by MALDI-MS

9.4.1 Study design for plasma/serum *N*-glycome analysis

This thesis reports on the glycomics measurement of several large cohorts (**Chapters 6**, **7**, and **8**), which, together with the literature, provides the opportunity to make some overarching statements. As it seems, findings that are presented in glycomics studies are typically small in effect size, and associate, next to the outcome of interest, as well with a broad range of other phenomena. For instance, a multitude of (inflammatory) disorders is marked by the galactosylation behavior of IgG-Fc[95,118-120], and many of the acute phase protein glycosylation behaviors are expected to be representative of overall mechanisms of acute inflammation rather than something that is disease-specific[116,121,122].

To find disease-specific biomarkers (and markers for stratification of patients) within the total plasma *N*-glycome, it makes sense to always include in the study the clinical information on age, sex, BMI and a marker for inflammation such as the level of C-reactive protein (CRP; or interleukin-6, but in our study CRP showed to be the better predictor) (**Chapter 6**). Without these measures it will be very hard to provide sensible interpretation of new clinical associations, and the specificity of new biomarkers would be limited as a consequence. Interestingly, while we have performed the largest mass spectrometric inquiry between clinical health markers and plasma *N*-glycosylation to date (**Chapter 6**), the majority of the glycomics variance is still not explained by clinical factors. This suggests

plenty of room still for future research into the cause and effect of glycan profiles, for example into the direction of genetic and epigenetic regulation[123-125]. One potential aspect of (epi)genetic regulation is the repeated observation that glycosylation is more variable between individuals than within individuals (in homeostasis) (**Chapter 7** and **8**)[51,126]. For the clinical exploitation of glycomics this means that a disease biomarker could preferentially lie in the repeat measurement of the glycans of an individual throughout their lifetime. A deviation from a standard profile of a person will then provide the warning call to proceed to more specific modes of diagnosis. In this possible future, MALDI-MS will have the benefit of throughput (**Chapter 8**)[5] – a useful feature for the analytical workload of population screening – but it remains to be seen if the relevant glycosylation characteristics can be detected by MS.

One example of the benefit of longitudinal analysis can already be found in this thesis, where we demonstrate the rather drastic upregulation of α2,3-linked sialylation with pregnancy on an individual level (**Chapter 7**). However, in general still the most basic of longitudinal analyses need to be performed with regard to glycomics, encompassing intraday variation (*e.g.*, dietary intake, tissue damage, acute inflammation, physical or mental activity), short-term interday variation (*e.g.*, menstrual cycle, virus infection), and the longer time-frames which have the most clinical relevance.

9.4.2 Functional studies

The analysis of plasma or serum *N*-glycosylation (**Chapters 5**, **6**, **7** and **8**) is a fundamentally phenotypical study. Whereas genetics can be presumed to be upstream to many diseases (a clear genetic defect leaves little speculation on causality), glycans observed in plasma or even single proteins are the end product of a large series of biological processes, making it challenging to determine cause and consequence. Factors influencing glycosylation are incredibly encompassing, from the abundance of glycosyltransferases and nucleotide sugars, *via* the time spent in the ER and Golgi, to the activity of intra- and extracellular recycling mechanisms[3,127,128].

Much of the functional research in glycomics, the targeted manipulation of glycosylation properties and observing their effects, has been performed on IgG[129-131]. It is an attractive object of study, because the protein is involved in several immunological processes (adaptive immune response and complement activation), and has altered function depending on its glycosylation[77,129-131]. In addition, pharmaceutical companies are interested in the modulation of IgG glycosylation to, *e.g.*, decrease toxicity, increase half-life, and increase ADCC[89,129,132]. However, functional studies outside of IgG have been relatively scarce.

One factor that is complicating functional studies is the translation from cell lines and animal models to the human situation, as glycosylation appears to be species-specific and differs

from one cell line to the next[3,18,107,108,133]. Mouse *N*-glycosylation, for which we have studied total plasma and IgG-Fc in healthy individuals (**Chapter 5**)[19], differs in sever ways from that of humans. For one, the mouse expresses predominantly *N*-glycolylneuraminic acid (as opposed to *N*-acetylneuraminic acid in humans)[3,18], and its *N*-glycans may carry α-linked galactoses (which are normally absent in humans)[20,134]. Further differences include a partial acetylation of sialic acids, the presence of GlcNAc-linked sialic acids for the triantennary species (possibly at the expense of the antennary fucosylation commonly found in humans) , and a relatively low level of IgG-Fc type glycans in plasma (**Chapter 5**)[19,20]. Next to this, mouse strains show to differ in glycosylation properties, which is most prominently visible by the degree of fucosylation in female individuals (ranging from 10% to 50% depending on the strain) (**Chapter 5**)[19].

For cell line work there are similar concerns, as cell lines such as CHO, HEK, and others differ in their respective glycosylation[107,108,133,135]. Fortunately, several strategies have been published that can provide control over protein glycoforms by *in vitro* or *in vivo* manipulation[131,136]. Isolated proteins can be modified by glycosidases and glycosyltransferases in presence of the appropriate nucleotide sugar[131], and cell systems have been described that make use of genetic knock-outs and knock-ins, and transient expression of the appropriate genes[131,135-137].

9.5 Concluding remarks

To summarize, this thesis describes several contributions to the high-throughput analysis of complex glycan samples by MALDI-MS, covering strategies of chemical derivatization (**Chapter 3**), automation (**Chapter 4**), and computational solutions for data preprocessing and statistical analysis (**Chapter 4** and **6**). The functionality of the developed platforms was demonstrated by several applications, including the characterization of mouse-strain plasma *N*-glycosylation (**Chapter 5**), the profiling of principally healthy human donors (**Chapter 6**), and an assessment of the intraindividual changes in RA patients throughout pregnancy (**Chapter 7** and **8**). Lastly, the workflow developed in this thesis for MALDI-TOF-MS with derivatized sialic acid residues was compared with other contemporary methods for high-throughput glycan analysis (**Chapter 8**).

A next large step for MALDI-MS method development would lie in the inclusion of internal standardization methods, both to decrease analytical variability and to limit the mathematical co-dependence of glycan signals. Analytical method hyphenation, parallel running of orthogonal methods for the same samples, and/or further advances in glycan characterization by MS(/MS) are still required to obtain the remaining relevant structural information of glycans in high-throughput fashion. Clinical glycomics would benefit from the translation of plasma-wide observations to particular proteins or groups of proteins, which can then subsequently be isolated for in-depth analysis. Lastly, the advent of longitudinal

glycomics studies is expected to be of considerable benefit to the translation of analytical methodology to clinical diagnostics and to facilitate patient stratification.

9.6 References

1 Apweiler, R., Hermjakob, H. & Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* **1473**, 4-8 (1999).

2 Khoury, G. A., Baliban, R. C. & Floudas, C. A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci Rep* **1**, doi:10.1038/srep00090 (2011).

3 Varki, A. *et al. Essentials of Glycobiology*. 2nd edn, (2009).

4 Harvey, D. J. Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates. *Mass Spectrom Rev* **18**, 349-450, doi:10.1002/(SICI)1098-2787(1999)18:6<349::AID-MAS1>3.0.CO;2-H (1999).

5 Huffman, J. E. *et al.* Comparative performance of four methods for high-throughput glycosylation analysis of immunoglobulin G in genetic and epidemiological research. *Mol Cell Proteomics* **13**, 1598-1610, doi:10.1074/mcp.M113.037465 (2014).

6 de Hoffmann, E. & Stroobant, V. *Mass spectrometry: principles and applications*. 3rd edn, (2007).

7 Borelli, V. *et al.* Plasma N-Glycome Signature of Down Syndrome. *J Proteome Res* **14**, 4232-4245, doi:10.1021/acs.jproteome.5b00356 (2015).

8 Bondt, A. *et al.* Immunoglobulin G (IgG) Fab glycosylation analysis using a new mass spectrometric high-throughput profiling method reveals pregnancy-associated changes. *Mol Cell Proteomics* **13**, 3029-3039, doi:10.1074/mcp.M114.039537 (2014).

9 Jansen, B. C. *et al.* Pregnancy-associated serum N-glycome changes studied by high-throughput MALDI-TOF-MS. *Sci Rep* **6**, 23296, doi:10.1038/srep23296 (2016).

10 Jansen, B. C. *et al.* MALDI-TOF-MS reveals differential N-linked plasma- and IgG-glycosylation profiles between mothers and their newborns. *Sci Rep* **6**, 34001, doi:10.1038/srep34001 (2016).

11 Rouwendal, G. J. *et al.* A comparison of anti-HER2 IgA and IgG1 in vivo efficacy is facilitated by high N-glycan sialylation of the IgA. *MAbs* **8**, 74-86, doi:10.1080/19420862.2015.1102812 (2016).

12 Meyer, S. *et al.* Improved in vivo anti-tumor effects of IgA-Her2 antibodies through half-life extension and serum exposure enhancement by FcRn targeting. *MAbs* **8**, 87-98, doi:10.1080/19420862.2015.1106658 (2016).

13 Reusch, D. *et al.* Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles--part 1: separation-based methods. *MAbs* **7**, 167-179, doi:10.4161/19420862.2014.986000 (2015).

14 Reusch, D. *et al.* Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles-Part 2: Mass spectrometric methods. *MAbs* **7**, 732-742, doi:10.1080/19420862.2015.1045173 (2015).

15 Shubhakar, A. *et al.* High-Throughput Analysis and Automation for Glycomics Studies. *Chromatographia* **78**, 321-333, doi:10.1007/s10337-014-2803-9 (2015).

16 Reiding, K. R., Blank, D., Kuijper, D. M., Deelder, A. M. & Wuhrer, M. High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification. *Anal Chem* **86**, 5784-5793, doi:10.1021/ac500335t (2014).

17 Galili, U., Shohet, S. B., Kobrin, E., Stults, C. L. & Macher, B. A. Man, apes, and Old World monkeys differ from other mammals in the expression of alpha-galactosyl epitopes on nucleated cells. *J Biol Chem* **263**, 17755-17762 (1988).

18 Varki, A. Diversity in the sialic acids. *Glycobiology* **2**, 25-40 (1992).

19 Reiding, K. R., Hipgrave Ederveen, A. L., Rombouts, Y. & Wuhrer, M. Murine Plasma N-Glycosylation Traits Associated with Sex and Strain. *J Proteome Res* **15**, 3489-3499, doi:10.1021/acs.jproteome.6b00071 (2016).
20 de Haan, N. *et al.* The N-Glycosylation of Mouse Immunoglobulin G (IgG)-Fragment Crystallizable Differs Between IgG Subclasses and Strains. *Front Immunol* **8**, 608, doi:10.3389/fimmu.2017.00608 (2017).
21 Reiding, K. R., Lonardi, E., Hipgrave Ederveen, A. L. & Wuhrer, M. Ethyl Esterification for MALDI-MS Analysis of Protein Glycosylation. *Methods Mol Biol* **1394**, 151-162, doi:10.1007/978-1-4939-3341-9_11 (2016).
22 de Haan, N. *et al.* Linkage-specific sialic acid derivatization for MALDI-TOF-MS profiling of IgG glycopeptides. *Anal Chem* **87**, 8284-8291, doi:10.1021/acs.analchem.5b02426 (2015).
23 Gomes de Oliveira, A. G. *et al.* A systematic study of glycopeptide esterification for the semi-quantitative determination of sialylation in antibodies. *Rapid Commun Mass Spectrom* **29**, 1817-1826, doi:10.1002/rcm.7287 (2015).
24 de Haan, N., Reiding, K. R. & Wuhrer, M. Sialic Acid Derivatization for the Rapid Subclass- and Sialic Acid Linkage-Specific MALDI-TOF-MS Analysis of IgG Fc-Glycopeptides. *Methods Mol Biol* **1503**, 49-62, doi:10.1007/978-1-4939-6493-2_5 (2017).
25 Holst, S. *et al.* Linkage-Specific in Situ Sialic Acid Derivatization for N-Glycan Mass Spectrometry Imaging of Formalin-Fixed Paraffin-Embedded Tissues. *Anal Chem* **88**, 5904-5913, doi:10.1021/acs.analchem.6b00819 (2016).
26 Wheeler, S. F., Domann, P. & Harvey, D. J. Derivatization of sialic acids for stabilization in matrix-assisted laser desorption/ionization mass spectrometry and concomitant differentiation of alpha(2 --> 3)- and alpha(2 --> 6)-isomers. *Rapid Commun Mass Spectrom* **23**, 303-312, doi:10.1002/rcm.3867 (2009).
27 Li, H., Gao, W., Feng, X., Liu, B. F. & Liu, X. MALDI-MS analysis of sialylated N-glycan linkage isomers using solid-phase two step derivatization method. *Anal Chim Acta* **924**, 77-85, doi:10.1016/j.aca.2016.04.023 (2016).
28 Nishikaze, T. *et al.* Differentiation of Sialyl Linkage Isomers by One-Pot Sialic Acid Derivatization for Mass Spectrometry-Based Glycan Profiling. *Anal Chem* **89**, 2353-2360, doi:10.1021/acs.analchem.6b04150 (2017).
29 Bladergroen, M. R. *et al.* Automation of High-Throughput Mass Spectrometry-Based Plasma N-Glycome Analysis with Linkage-Specific Sialic Acid Esterification. *J Proteome Res* **14**, 4080-4086, doi:10.1021/acs.jproteome.5b00538 (2015).
30 Jansen, B. C. *et al.* MassyTools: A High-Throughput Targeted Data Processing Tool for Relative Quantitation and Quality Control Developed for Glycomic and Glycoproteomic MALDI-MS. *J Proteome Res* **14**, 5088-5098, doi:10.1021/acs.jproteome.5b00658 (2015).
31 Reiding, K. R. *et al.* Human Plasma N-glycosylation as Analyzed by Matrix-Assisted Laser Desorption/Ionization-Fourier Transform Ion Cyclotron Resonance-MS Associates with Markers of Inflammation and Metabolic Health. *Mol Cell Proteomics* **16**, 228-242, doi:10.1074/mcp.M116.065250 (2017).
32 Gong, B. *et al.* N-glycosylamine-mediated isotope labeling for mass spectrometry-based quantitative analysis of N-linked glycans. *Anal Bioanal Chem* **405**, 5825-5831, doi:10.1007/s00216-013-6988-9 (2013).
33 Lauber, M. A. *et al.* Rapid Preparation of Released N-Glycans for HILIC Analysis Using a Labeling Reagent that Facilitates Sensitive Fluorescence and ESI-MS Detection. *Anal Chem* **87**, 5401-5409, doi:10.1021/acs.analchem.5b00758 (2015).
34 Zaia, J. Mass Spectrometry and Glycomics. *OMICS : a Journal of Integrative Biology* **14**, 401-418, doi:10.1089/omi.2009.0146 (2010).
35 Morelle, W. & Michalski, J. C. Analysis of protein glycosylation by mass spectrometry. *Nat Protoc* **2**, 1585-1602, doi:10.1038/nprot.2007.227 (2007).

36 Ruhaak, L. R., Steenvoorden, E., Koeleman, C. A. M., Deelder, A. M. & Wuhrer, M. 2-Picoline-borane: A non-toxic reducing agent for oligosaccharide labeling by reductive amination. *Proteomics* **10**, 2330-2336, doi:10.1002/pmic.200900804 (2010).
37 Ruhaak, L. R. *et al.* Hydrophilic interaction chromatography-based high-throughput sample preparation method for N-glycan analysis from total human plasma glycoproteins. *Anal Chem* **80**, 6119-6126, doi:10.1021/ac800630x (2008).
38 Selman, M. H. J., Hemayatkar, M., Deelder, A. M. & Wuhrer, M. Cotton HILIC SPE Microtips for Microscale Purification and Enrichment of Glycans and Glycopeptides. *Anal Chem* **83**, 2492-2499, doi:10.1021/ac1027116 (2011).
39 Yang, S., Hu, Y., Sokoll, L. & Zhang, H. Simultaneous quantification of N- and O-glycans using a solid-phase method. *Nat Protoc* **12**, 1229-1244, doi:10.1038/nprot.2017.034 (2017).
40 Zhou, H. *et al.* Universal Solid-Phase Reversible Sample-Prep for Concurrent Proteome and N-Glycome Characterization. *J Proteome Res* **15**, 891-899, doi:10.1021/acs.jproteome.5b00865 (2016).
41 Miura, Y. *et al.* BlotGlycoABCTM, an integrated glycoblotting technique for rapid and large scale clinical glycomics. *Mol Cell Proteomics* **7**, 370-377, doi:10.1074/mcp.M700377-MCP200 (2008).
42 Prien, J. M., Prater, B. D., Qin, Q. & Cockrill, S. L. Mass Spectrometric-Based Stable Isotopic 2-Aminobenzoic Acid Glycan Mapping for Rapid Glycan Screening of Biotherapeutics. *Anal Chem* **82**, 1498-1508, doi:10.1021/ac902617t (2010).
43 Walker, S. H., Budhathoki-Uprety, J., Novak, B. M. & Muddiman, D. C. Stable-Isotope Labeled Hydrophobic Hydrazide Reagents for the Relative Quantification of N-Linked Glycans by Electrospray Ionization Mass Spectrometry. *Anal Chem* **83**, 6738-6745, doi:10.1021/ac201376q (2011).
44 Zhang, P. *et al.* Relative quantitation of glycans using stable isotopic labels 1-(d0/d5) phenyl-3-methyl-5-pyrazolone by mass spectrometry. *Anal Biochem* **418**, 1-9, doi:10.1016/j.ab.2011.07.006 (2011).
45 Atwood, J. A., 3rd *et al.* Quantitation by isobaric labeling: applications to glycomics. *J Proteome Res* **7**, 367-374, doi:10.1021/pr070476i (2008).
46 Shah, P. *et al.* Mass Spectrometric Analysis of Sialylated Glycans Using Solid Phase Labeling of Sialc Acids. *Anal Chem* **85**, 3606-3613, doi:10.1021/ac3033867 (2013).
47 Klein, A. Human total serum N-glycome. *Adv Clin Chem* **46**, 51-85 (2008).
48 Clerc, F. *et al.* Human plasma protein N-glycosylation. *Glycoconj J* **33**, 309-343, doi:10.1007/s10719-015-9626-2 (2016).
49 Kolarich, D., Weber, A., Turecek, P. L., Schwarz, H. P. & Altmann, F. Comprehensive glyco-proteomic analysis of human alpha1-antitrypsin and its charge isoforms. *Proteomics* **6**, 3369-3380, doi:10.1002/pmic.200500751 (2006).
50 Zhang, L., Luo, S. & Zhang, B. The use of lectin microarray for assessing glycosylation of therapeutic proteins. *MAbs* **8**, 524-535, doi:10.1080/19420862.2016.1149662 (2016).
51 Hennig, R. *et al.* Towards personalized diagnostics via longitudinal study of the human plasma N-glycome. *Biochim Biophys Acta* **1860**, 1728-1738, doi:10.1016/j.bbagen.2016.03.035 (2016).
52 Saldova, R. *et al.* Association of N-glycosylation with breast carcinoma and systemic features using high-resolution quantitative UPLC. *J Proteome Res* **13**, 2314-2327, doi:10.1021/pr401092y (2014).
53 Harvey, D. J. Fragmentation of Negative Ions from Carbohydrates: Part 1. Use of Nitrate and Other Anionic Adducts for the Production of Negative Ion Electrospray Spectra from N-linked Carbohydrates. *J Am Soc Mass Spectr* **16**, 622-630, doi:https://doi.org/10.1016/j.jasms.2005.01.004 (2005).

54 Mechref, Y., Kang, P. & Novotny, M. V. Differentiating structural isomers of sialylated glycans by matrix-assisted laser desorption/ionization time-of-flight/time-of-flight tandem mass spectrometry. *Rapid Commun Mass Spectrom* **20**, 1381-1389, doi:10.1002/rcm.2445 (2006).
55 Stumpo, K. A. & Reinhold, V. N. The N-glycome of human plasma. *J Proteome Res* **9**, 4823-4830, doi:10.1021/pr100528k (2010).
56 Halim, A. *et al.* Assignment of saccharide identities through analysis of oxonium ion fragmentation profiles in LC-MS/MS of glycopeptides. *J Proteome Res* **13**, 6024-6032, doi:10.1021/pr500898r (2014).
57 Zhou, S., Dong, X., Veillon, L., Huang, Y. & Mechref, Y. LC-MS/MS analysis of permethylated N-glycans facilitating isomeric characterization. *Anal Bioanal Chem* **409**, 453-466, doi:10.1007/s00216-016-9996-8 (2017).
58 Zhong, X. *et al.* Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry for Quantitative Analysis of Glycans Labeled with Multiplex Carbonyl-Reactive Tandem Mass Tags. *Anal Chem* **87**, 6527-6534, doi:10.1021/acs.analchem.5b01835 (2015).
59 Guttman, M. & Lee, K. K. Site-Specific Mapping of Sialic Acid Linkage Isomers by Ion Mobility Spectrometry. *Anal Chem* **88**, 5212-5217, doi:10.1021/acs.analchem.6b00265 (2016).
60 Selman, M. H. J. *et al.* Immunoglobulin G Glycopeptide Profiling by Matrix-Assisted Laser Desorption Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal Chem* **82**, 1073-1081, doi:10.1021/ac9024413 (2010).
61 Bondt, A. *et al.* Longitudinal monitoring of immunoglobulin A glycosylation during pregnancy by simultaneous MALDI-FTICR-MS analysis of N- and O-glycopeptides. *Sci Rep* **6**, 27955, doi:10.1038/srep27955 (2016).
62 Moh, E. S., Lin, C. H., Thaysen-Andersen, M. & Packer, N. H. Site-Specific N-Glycosylation of Recombinant Pentameric and Hexameric Human IgM. *J Am Soc Mass Spectr* **27**, 1143-1155, doi:10.1007/s13361-016-1378-0 (2016).
63 Hinneburg, H. *et al.* The Art of Destruction: Optimizing Collision Energies in Quadrupole-Time of Flight (Q-TOF) Instruments for Glycopeptide-Based Glycoproteomics. *J Am Soc Mass Spectr* **27**, 507-519, doi:10.1007/s13361-015-1308-6 (2016).
64 Selman, M. H. J. *et al.* Fc specific IgG glycosylation profiling by robust nano-reverse phase HPLC-MS using a sheath-flow ESI sprayer interface. *J Proteomics* **75**, 1318-1329, doi:10.1016/j.jprot.2011.11.003 (2012).
65 Wuhrer, M., Catalina, M. I., Deelder, A. M. & Hokke, C. H. Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J Chromatogr B Analyt Technol Biomed Life Sci* **849**, 115-128, doi:10.1016/j.jchromb.2006.09.041 (2007).
66 Dodds, E. D. Gas-phase dissociation of glycosylated peptide ions. *Mass Spectrom Rev* **31**, 666-682, doi:10.1002/mas.21344 (2012).
67 Alley, W. R., Jr., Mechref, Y. & Novotny, M. V. Characterization of glycopeptides by combining collision-induced dissociation and electron-transfer dissociation mass spectrometry data. *Rapid Commun Mass Spectrom* **23**, 161-170, doi:10.1002/rcm.3850 (2009).
68 Nilsson, J. Liquid chromatography-tandem mass spectrometry-based fragmentation analysis of glycopeptides. *Glycoconj J* **33**, 261-272, doi:10.1007/s10719-016-9649-3 (2016).
69 Nishikaze, T., Kawabata, S. & Tanaka, K. Fragmentation characteristics of deprotonated N-linked glycopeptides: influences of amino acid composition and sequence. *J Am Soc Mass Spectr* **25**, 988-998, doi:10.1007/s13361-014-0854-7 (2014).
70 Nishikaze, T., Nakamura, T., Jinmei, H. & Amano, J. Negative-ion MALDI-MS2 for discrimination of α2,3- and α2,6-sialylation on glycopeptides labeled with a pyrene derivative. *Journal of Chromatography B* **879**, 1419-1428, doi:http://dx.doi.org/10.1016/j.jchromb.2010.10.032 (2011).
71 Nishikaze, T., Kawabata, S. & Tanaka, K. In-Depth Structural Characterization of N-Linked Glycopeptides Using Complete Derivatization for Carboxyl Groups Followed by Positive- and

Negative-Ion Tandem Mass Spectrometry. *Anal Chem* **86**, 5360-5369, doi:10.1021/ac500340t (2014).

72 Bunkenborg, J., Pilch, B. J., Podtelejnikov, A. V. & Wisniewski, J. R. Screening for N-glycosylated proteins by liquid chromatography mass spectrometry. *Proteomics* **4**, 454-465, doi:10.1002/pmic.200300556 (2004).

73 Liu, T. *et al.* Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. *J Proteome Res* **4**, 2070-2080, doi:10.1021/pr0502065 (2005).

74 Jia, W. *et al.* A strategy for precise and large scale identification of core fucosylated glycoproteins. *Mol Cell Proteomics* **8**, 913-923, doi:10.1074/mcp.M800504-MCP200 (2009).

75 Warrack, B. M. *et al.* Normalization strategies for metabonomic analysis of urine samples. *J Chromatogr B Analyt Technol Biomed Life Sci* **877**, 547-552, doi:10.1016/j.jchromb.2009.01.007 (2009).

76 Wu, Y. & Li, L. Sample normalization methods in quantitative metabolomics. *Journal of chromatography. A* **1430**, 80-95, doi:10.1016/j.chroma.2015.12.007 (2016).

77 Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M. & Dwek, R. A. The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu Rev Immunol* **25**, 21-50, doi:10.1146/annurev.immunol.25.022106.141702 (2007).

78 Zauner, G. *et al.* Glycoproteomic analysis of antibodies. *Mol Cell Proteomics* **12**, 856-865, doi:10.1074/mcp.R112.026005 (2013).

79 Nairn, A. V. *et al.* Regulation of glycan structures in animal tissues: transcript profiling of glycan-related genes. *J Biol Chem* **283**, 17298-17313, doi:10.1074/jbc.M801964200 (2008).

80 Freeze, H. H. Genetic defects in the human glycome. *Nat Rev Genet* **7**, 537-551, doi:10.1038/nrg1894 (2006).

81 Bondt, A. *et al.* Association between galactosylation of immunoglobulin G and improvement of rheumatoid arthritis during pregnancy is independent of sialylation. *J Proteome Res* **12**, 4522-4531, doi:10.1021/pr400589m (2013).

82 Plomp, R. *et al.* Subclass-specific IgG glycosylation is associated with markers of inflammation and metabolic health. *Sci Rep* **7**, 12325, doi:10.1038/s41598-017-12495-0 (2017).

83 Kristic, J. *et al.* Glycans are a novel biomarker of chronological and biological ages. *J Gerontol A Biol Sci Med Sci* **69**, 779-789, doi:10.1093/gerona/glt190 (2014).

84 Trbojevic Akmacic, I. *et al.* Inflammatory bowel disease associates with proinflammatory potential of the immunoglobulin G glycome. *Inflamm Bowel Dis* **21**, 1237-1247, doi:10.1097/MIB.0000000000000372 (2015).

85 Borghesi, L. & Milcarek, C. From B cell to plasma cell: regulation of V(D)J recombination and antibody secretion. *Immunol Res* **36**, 27-32, doi:10.1385/IR:36:1:27 (2006).

86 Arnold, J. N. *et al.* Human serum IgM glycosylation - Identification of glycoforms that can bind to mannan-binding lectin. *Journal of Biological Chemistry* **280**, 29080-29087, doi:10.1074/jbc.M504528200 (2005).

87 Mattu, T. S. *et al.* The glycosylation and structure of human serum IgA1, Fab, and Fc regions and the role of N-glycosylation on Fcalpha receptor interactions. *J Biol Chem* **273**, 2260-2272 (1998).

88 Wang, J. *et al.* Fc-glycosylation of IgG1 is modulated by B-cell stimuli. *Mol Cell Proteomics* **10**, M110 004655, doi:10.1074/mcp.M110.004655 (2011).

89 Reusch, D. & Tejada, M. L. Fc glycans of therapeutic antibodies as critical quality attributes. *Glycobiology* **25**, 1325-1334, doi:10.1093/glycob/cwv065 (2015).

90 Wormald, M. R. *et al.* Variations in oligosaccharide-protein interactions in immunoglobulin G determine the site-specific glycosylation profiles and modulate the dynamic motion of the Fc oligosaccharides. *Biochemistry-Us* **36**, 1370-1380, doi:10.1021/bi9621472 (1997).

91 Jefferis, R. Glycosylation as a strategy to improve antibody-based therapeutics. *Nature reviews. Drug discovery* **8**, 226-234, doi:10.1038/nrd2804 (2009).
92 Raymond, C. l. *et al.* in *Glycosylation* (ed Stefana Petrescu) Ch. 17 (InTech, 2012).
93 Wormald, M. R. *et al.* Variations in oligosaccharide-protein interactions in immunoglobulin G determine the site-specific glycosylation profiles and modulate the dynamic motion of the Fc oligosaccharides. *Biochemistry-Us* **36**, 1370-1380, doi:Doi 10.1021/Bi9621472 (1997).
94 Yamaguchi, Y. *et al.* Dynamics of the carbohydrate chains attached to the Fc portion of immunoglobulin G as studied by NMR spectroscopy assisted by selective C-13 labeling of the glycans. *J Biomol NMR* **12**, 385-394, doi:Doi 10.1023/A:1008392229694 (1998).
95 Dall'Olio, F. *et al.* N-glycomic biomarkers of biological aging and longevity: a link with inflammaging. *Ageing Res Rev* **12**, 685-698, doi:10.1016/j.arr.2012.02.002 (2013).
96 Catera, M. *et al.* Identification of novel plasma glycosylation-associated markers of aging. *Oncotarget* **7**, 7455-7468, doi:10.18632/oncotarget.7059 (2016).
97 Van Slyke, D. D. *et al.* THE ESTIMATION OF PLASMA PROTEIN CONCENTRATION FROM PLASMA SPECIFIC GRAVITY. *Journal of Biological Chemistry* **183**, 331-347 (1950).
98 Adamczyk, B., Struwe, W. B., Ercan, A., Nigrovic, P. A. & Rudd, P. M. Characterization of Fibrinogen Glycosylation and Its Importance for Serum/Plasma N-Glycome Analysis. *J Proteome Res* **12**, 444-454, doi:10.1021/pr300813h (2013).
99 Fu, D. & van Halbeek, H. N-glycosylation site mapping of human serotransferrin by serial lectin affinity chromatography, fast atom bombardment-mass spectrometry, and 1H nuclear magnetic resonance spectroscopy. *Anal Biochem* **206**, 53-63 (1992).
100 Pompach, P. *et al.* Site-specific glycoforms of haptoglobin in liver cirrhosis and hepatocellular carcinoma. *Mol Cell Proteomics* **12**, 1281-1293, doi:10.1074/mcp.M112.023259 (2013).
101 Imre, T. *et al.* Mass spectrometric and linear discriminant analysis of N-glycans of human serum alpha-1-acid glycoprotein in cancer patients and healthy individuals. *J Proteomics* **71**, 186-197, doi:10.1016/j.jprot.2008.04.005 (2008).
102 Olofsson, S. O. *et al.* Apolipoprotein B: structure, biosynthesis and role in the lipoprotein assembly process. *Atherosclerosis* **68**, 1-17 (1987).
103 Olofsson, S. O. & Boren, J. Apolipoprotein B secretory regulation by degradation. *Arterioscler Thromb Vasc Biol* **32**, 1334-1338, doi:10.1161/ATVBAHA.112.251116 (2012).
104 Harazono, A., Kawasaki, N., Kawanishi, T. & Hayakawa, T. Site-specific glycosylation analysis of human apolipoprotein B100 using LC/ESI MS/MS. *Glycobiology* **15**, 447-462, doi:10.1093/glycob/cwi033 (2005).
105 Lottspeich, F., Kellermann, J., Henschen, A., Foertsch, B. & Muller-Esterl, W. The amino acid sequence of the light chain of human high-molecular-mass kininogen. *Eur J Biochem* **152**, 307-314 (1985).
106 Poon, I. K., Patel, K. K., Davis, D. S., Parish, C. R. & Hulett, M. D. Histidine-rich glycoprotein: the Swiss Army knife of mammalian plasma. *Blood* **117**, 2093-2101, doi:10.1182/blood-2010-09-303842 (2011).
107 Croset, A. *et al.* Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells. *J Biotechnol* **161**, 336-348, doi:10.1016/j.jbiotec.2012.06.038 (2012).
108 Bohm, E. *et al.* Expression of recombinant human coagulation factors VII (rFVII) and IX (rFIX) in various cell types, glycosylation analysis, and pharmacokinetic comparison. *BMC Proc* **5 Suppl 8**, P23, doi:10.1186/1753-6561-5-S8-P23 (2011).
109 West, M. B. *et al.* Analysis of Site-specific Glycosylation of Renal and Hepatic γ-Glutamyl Transpeptidase from Normal Human Tissue. *Journal of Biological Chemistry* **285**, 29511-29524 (2010).
110 Ikegami, K. *et al.* Tissue-specific posttranslational modification allows functional targeting of thyrotropin. *Cell Rep* **9**, 801-810, doi:10.1016/j.celrep.2014.10.006 (2014).

111 Gramolini, A., Lau, E. & Liu, P. P. Identifying Low-Abundance Biomarkers. *Circulation* **134**, 286 (2016).
112 Ruhaak, L. R. *et al.* Targeted biomarker discovery by high throughput glycosylation profiling of human plasma alpha1-antitrypsin and immunoglobulin A. *PLoS One* **8**, e73082, doi:10.1371/journal.pone.0073082 (2013).
113 Ruhaak, L. R. *et al.* Plasma protein N-glycan profiles are associated with calendar age, familial longevity and health. *J Proteome Res* **10**, 1667-1674, doi:10.1021/pr1009959 (2011).
114 Ercan, A. *et al.* Estrogens regulate glycosylation of IgG in women and men. *JCI Insight* **2**, e89703, doi:10.1172/jci.insight.89703 (2017).
115 Brinkman-Van der Linden, C. M. *et al.* Oral estrogen treatment induces a decrease in expression of sialyl Lewis x on alpha 1-acid glycoprotein in females and male-to-female transsexuals. *Glycobiology* **6**, 407-412 (1996).
116 Albrecht, S., Unwin, L., Muniyappa, M. & Rudd, P. M. Glycosylation as a marker for inflammatory arthritis. *Cancer Biomark* **14**, 17-28, doi:10.3233/CBM-130373 (2014).
117 Ceciliani, F. & Pocacqua, V. The acute phase protein alpha1-acid glycoprotein: a model for altered glycosylation during diseases. *Curr Protein Pept Sci* **8**, 91-108 (2007).
118 Theodoratou, E. *et al.* Glycosylation of plasma IgG in colorectal cancer prognosis. *Sci Rep* **6**, 28098, doi:10.1038/srep28098 (2016).
119 Miyoshi, E. *et al.* Role of aberrant IgG glycosylation in the pathogenesis of inflammatory bowel disease. *Proteomics Clin Appl* **10**, 384-390, doi:10.1002/prca.201500089 (2016).
120 Lemmers, R. F. H. *et al.* IgG glycan patterns are associated with type 2 diabetes in independent European populations. *Biochim Biophys Acta* **1861**, 2240-2249, doi:10.1016/j.bbagen.2017.06.020 (2017).
121 McCarthy, C. *et al.* The role and importance of glycosylation of acute phase proteins with focus on alpha-1 antitrypsin in acute and chronic inflammatory conditions. *J Proteome Res* **13**, 3131-3143, doi:10.1021/pr500146y (2014).
122 Peracaula, R., Sarrats, A. & Rudd, P. M. Liver proteins as sensor of human malignancies and inflammation. *Proteomics Clin Appl* **4**, 426-431, doi:10.1002/prca.200900170 (2010).
123 Huffman, J. E. *et al.* Polymorphisms in B3GAT1, SLC9A9 and MGAT5 are associated with variation within the human plasma N-glycome of 3533 European adults. *Hum Mol Genet* **20**, 5000-5011, doi:10.1093/hmg/ddr414 (2011).
124 Zoldos, V. *et al.* Epigenetic silencing of HNF1A associates with changes in the composition of the human plasma N-glycome. *Epigenetics-Us* **7**, 164-172, doi:10.4161/epi.7.2.18918 (2012).
125 Menni, C. *et al.* Glycosylation of immunoglobulin g: role of genetic and epigenetic influences. *PLoS One* **8**, e82558, doi:10.1371/journal.pone.0082558 (2013).
126 Gornik, O. *et al.* Stability of N-glycan profiles in human plasma. *Glycobiology* **19**, 1547-1553, doi:10.1093/glycob/cwp134 (2009).
127 Moremen, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol* **13**, 448-462, doi:10.1038/nrm3383 (2012).
128 Yang, W. H. *et al.* An intrinsic mechanism of secreted protein aging and turnover. *Proc Natl Acad Sci U S A* **112**, 13657-13662, doi:10.1073/pnas.1515464112 (2015).
129 Thomann, M. *et al.* In vitro glycoengineering of IgG1 and its effect on Fc receptor binding and ADCC activity. *PLoS One* **10**, e0134949, doi:10.1371/journal.pone.0134949 (2015).
130 Dalziel, M., Crispin, M., Scanlan, C. N., Zitzmann, N. & Dwek, R. A. Emerging principles for the therapeutic exploitation of glycosylation. *Science* **343**, 1235681, doi:10.1126/science.1235681 (2014).
131 Dekkers, G. *et al.* Decoding the Human Immunoglobulin G-Glycan Repertoire Reveals a Spectrum of Fc-Receptor- and Complement-Mediated-Effector Activities. *Front Immunol* **8**, 877, doi:10.3389/fimmu.2017.00877 (2017).

132 Thomann, M., Reckermann, K., Reusch, D., Prasser, J. & Tejada, M. L. Fc-galactosylation modulates antibody-dependent cellular cytotoxicity of therapeutic antibodies. *Mol Immunol* **73**, 69-75, doi:10.1016/j.molimm.2016.03.002 (2016).

133 Holst, S. *et al.* N-glycosylation Profiling of Colorectal Cancer Cell Lines Reveals Association of Fucosylation with Differentiation and Caudal Type Homebox 1 (CDX1)/Villin mRNA Expression. *Mol Cell Proteomics* **15**, 124-140, doi:10.1074/mcp.M115.051235 (2016).

134 Chung, C. H. *et al.* Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. *New Engl J Med* **358**, 1109-1117, doi:Doi 10.1056/Nejmoa074943 (2008).

135 Yang, Q. *et al.* Glycan Remodeling of Human Erythropoietin (EPO) Through Combined Mammalian Cell Engineering and Chemoenzymatic Transglycosylation. *ACS Chem Biol* **12**, 1665-1673, doi:10.1021/acschembio.7b00282 (2017).

136 Yang, Z. *et al.* Engineered CHO cells for production of diverse, homogeneous glycoproteins. *Nat Biotechnol* **33**, 842-844, doi:10.1038/nbt.3280 (2015).

137 Hamilton, S. R. *et al.* Production of Complex Human Glycoproteins in Yeast. *Science* **301**, 1244 (2003).