Glycomics based biomarkers of the rate of aging : development and applications of high-throughput N-glycan analysis

Ruhaak, L.R.

Chapter 1.
Introduction

Partially based on:

Glycan labeling strategies and their use in identification and quantification
L.R. Ruhaak, G. Zauner, C. Huhn, C. Bruggink, A.M. Deelder, M. Wuhrer
Anal. Bioanal. Chem. (2010) 397; 3457-81.

and

IgG glycosylation analysis
C. Huhn, M. H. J. Selman, L. R. Ruhaak, A. M. Deelder, M. Wuhrer
Proteomics (2009) 9; 882-913

Glycosylation is the enzymatic addition of oligosaccharides (also known as glycans) to proteins and lipids. More than 50% of human proteins [1] are glycosylated and proper glycosylation is essential for the survival of most organisms as the glycans have important functions in biological processes such as: protein folding [1], protein clearance [2], cell adhesion [3;4], receptor binding and receptor activation [5;6]. As glycans are often branched structures, where monosaccharides may be linked in several different ways, protein glycosylation may be very diverse. Moreover, it is a dynamic equilibrium: within an individual, the glycan signature is highly reproducible in a given physiological state [7;8], however, when the physiological state changes, e.g. due to aging or disease, the glycan pattern can change dramatically [7]. Therefore, protein glycosylation is an important group of potential biomarkers, as recently reviewed by Packer et al. [9]. Markers for longevity which reflect the health condition and predict healthy aging would be valuable in aging research. Such markers, which indicate biological age of individuals instead of calendar age, have so far hardly been identified. The studies in this thesis are focused on the analysis of protein glycosylation with an emphasis on the development of potential candidate markers for familial longevity, which might reflect healthy aging.

Human glycans are built from a limited number of monosaccharides, of which the most abundant ones are: glucose (Glc), galactose (Gal), mannose (Man), fucose (Fuc), N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc) and N-acetylneuraminic acid (NeuAc). These monosaccharides are linked together by glycosidic linkages to form oligosaccharides or glycans. Typical examples of glycosylation types are N-glycans, O-glycans, glycosphingolipids and proteoglycans
(Figure 1 1).

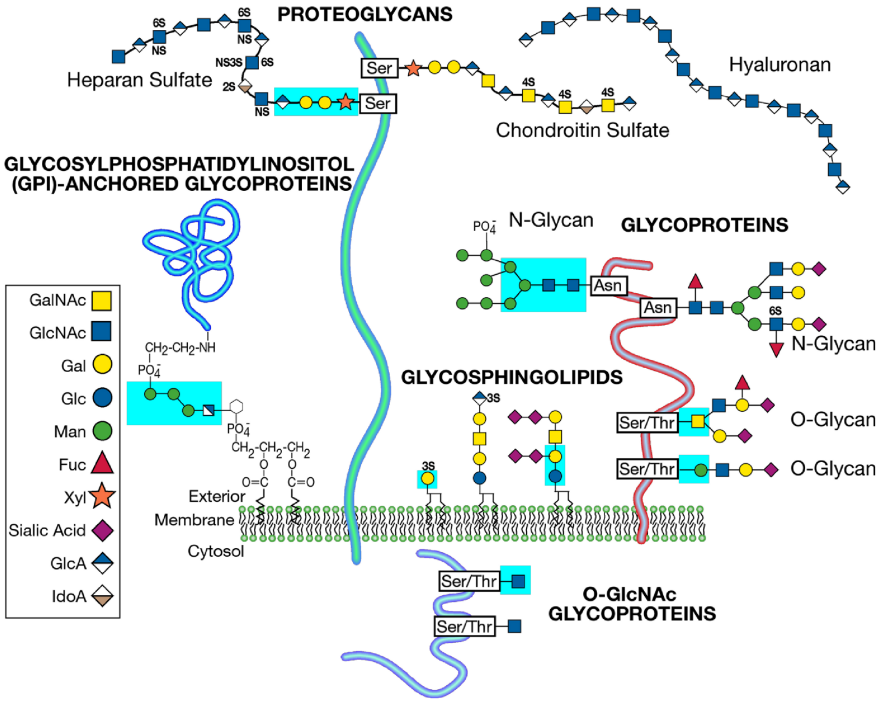

Figure 1-1. *Types of glycosylation (Reprinted from [10])*

N-glycans are covalently linked to the asparagine in the consensus sequence Asn-X-Ser/Thr and they always comprise a five-monosaccharide core consisting of two GlcNAcs and three mannoses. All studies in this thesis have been focused on N-glycans from human plasma proteins, therefore, other types of glycosylation will not be discussed further in this chapter/introduction. About 30% of the available N-glycosylation consensus sequons in the human proteome are actually occupied by N-glycans [11]. Moreover, some protein glycosylation sites may be found both in occupied and non-occupied form. On most glycosylation sites different glycan structures occur, thus giving rise to a heterogeneous set of variants for each glycoprotein. The analysis of the glycosylation of a protein therefore results in a profile representing the entire bouguet of the so-called glycoforms of the protein.

N-glycan biosynthesis

N-glycan biosynthesis, which is nowadays well understood [10-12], is a cellular process which takes place in the endoplasmatic reticulum and the golgi apparatus by several enzymes including glycosyltransferases and endoglycosidases. These enzymes are remarkably substrate specific, and can usually only use the product of the previous step in the pathway as their substrate [13], resulting in a robust, though non-plastic pathway.

The first step in N-glycan biosynthesis takes place on the cytoplasmic side of the endoplasmatic reticulum (ER) and comprises the addition of a phosphorylated N-acetylglucosamine to a dolichol phosphate substrate [12]. The then formed dolichol pyrophosphate N-acetylglucosamine is subsequently elongated at the N-acetylglucosamine with six additional monosaccharides (one GlcNAc and five mannoses), prior to translocation to the lumenal side of the ER. After addition of four more mannoses, the assembly of the precursor is completed with the linear addition of three glucose residues. The complete oligosaccharide is then transferred from the dolichol precursor to the asparagines in the Asn-X-Ser/Thr sequon, with the aid of a protein complex called oligosaccharyltransferase (OST). The chaperone proteins calnexin and calreticulin then facilitate the correct folding of the protein. When the protein is folded correctly, all three glucoses as well as one mannose are enzymatically removed from the glycoprotein prior to transfer to the Golgi apparatus for further processing by an extensive set of glycosidases and glycosyltransferases. Misfolded proteins are transferred to the proteasome for degradation.

Based on the effectiveness of the mannosidases in the biosynthetic pathway, which may now remove several of the the mannoses from the glycan, N-glycans can be subdivided in three groups, as depicted in Figure 1 2. They generally share the $GlcNAc_2Man_3$ core, which was generated in the ER, but the mannoses may be removed to different extents, resulting in three different patterns: 1. In high-mannose type N-glycans, the $GlcNAc_2Man_3$ core with the six mannoses is only slightly trimmed: one to three branches of maximal 2 mannoses remain on the $GlcNAc_2Man_3$

core. One chain is attached to the α1-3 branch, while the α1-6 linked mannose can be extended with two mannose branches. 2. Complex type N-glycans, where all six outer mannoses are removed. Each of the two external mannoses of the core is then extended with one or two GlcNAcs, which in turn are extended with one galactose and, possibly, a sialic acid (in humans N-acetylneuraminic acid or NeuAc). Complex type N-glycan may carry a bisecting GlcNAc, which is β1-4 linked to the β-linked core-mannose. 3. The hybrid type N-glycans can be regarded as a mixture of the two previous types: the mannoses of the α1-3 branch are all trimmed, while the several mannoses remain on the α1-6 arm. The α1-3 arm may be exteded with one or two complex type antenna.

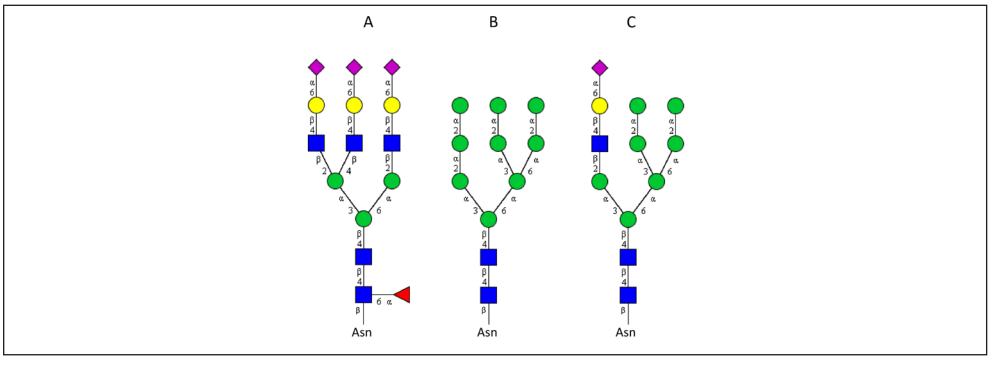

Figure 1-2. *Three classes of N-glycans. All N-glycans consist of a GlcNAc2Man3 core, which is then elongated according to the complex type (A), the high-mannose type (B) or the hybrid type (C). For symbol key see Figure 1 1.*

After biosynthesis, most glycoproteins are transported to the plasma membrane of the cell; however, some glycoproteins are excreted into the extracellular space, and can be monitored in blood plasma. Most of the excreted glycoproteins in human plasma are derived from either the liver cells or B-lymphocytes.

Altered protein glycosylation as potential biomarkers.

The diversity of the glycoform population is cell- and protein- specific. In a given physiological state, the glycan signature is highly reproducible [7;8], however, when the physiological state changes, e.g. due to aging or disease, the glycan pattern can

change dramatically (e.g [7]). Due to the large variability of glycosylation and the reflection of several physiological states in glycosylation patterns, protein glycosylation has been marked as a large pool of potential biomarkers and should be considered similar to genomics or proteomics [9].

To evaluate the potential of glycans or glycosylation patterns as biomarkers, it is necessary to analyse these features. As it is difficult to get access to cellular body materials, which is the original source of glycans, it is often decided to analyse body fluids, which may reflect the cellular processes. This approach is similar to proteomics and metabolomics, where the use of bodyfluids is widely applied (e.g. [14-16]). Glycosylation patterns have most often been studied in blood derived plasma or serum but also tissue specific glycosylation and glycosylation of tear fluid [17], saliva, cerebrospinal fluid [18] and urine have been studied.

Changes in glycosylation pattern have been recorded with increasing calendar age and for several disease states, and reviewing them all here would be too extensive, therefore we will focus on changes associated with calendar age and briefly discuss altered glycosylation patterns observed with some major age-related diseases: diabetes, cancer, Alzheimers disease and rheumatoid arthritis Reviews on altered glycosylation profiles with cancer states and liver diseased have recently been published ([7] and [19], respectively). This part is first focused on changes in glycosylation that were observed in total plasma, while a second part focuses on changes in altered glycosylation observed on specific glycoproteins.

Glycan-based markers in blood derived body fluids

Associations of total plasma protein glycosylation patterns with calendar age have been evaluated in a study population of 100 Belgian individuals, subdivided in five sex-matched groups of 20, 30, 40, 50 and 60 years of age [20]. As compared to subjects of 20 years of age, elderly individuals of ages above 50 had increased levels of non-galactosylated glycans, while the levels of galactosylated structures decreased with increasing calendar age. In the same study, a population of 120 Italian cente-

narians was compared to 79 elderly (mean age 81) and 63 middle-aged (mean age 44) individuals. In this high-age population, changes in plasma protein glycosylation which were similar to those observed for the Belgian population were observed as a function of age. This indicates that the changes in plasma protein glycosylation with age can be extrapolated to very high ages.

In the last two years, studies have been published on Croatian island populations regarding the heritability and variability of plasma N-glycosylation in healthy individuals [21], the genetic background of N-glycosylation [22;23] and the associations between chronological aging, smoke and lipid profiles [24].

In the first study [21], a cohort of 1008 individuals (415 men and 593 women) in the age of 18-93 was investigated. N-glycans were profiled using HPLC and 33 glycan features were obtained. Very large biological variability in glycosylation was observed, with coefficients of variation ranging from 6.4 to 50 %. Moreover a broad range of variation in heritability could be observed, ranging from insignificant to high. Therefore, it was hypothesized that some glycans are mostly controlled by genetic factors, wile other glycans are mainly under environmental control [21].

Genome-wide association data were obtained from 991 individuals from this study using Illumina Human Hap 300 [23]. As only one of the features in the N-glycan HPLC analysis could be identified as a single glycan ($Man_3GlcNAc_5$) and this glycan also showed substantial heritability, this glycan was evaluated for associations with single nucleotide polymorphisms (SNPs). The most significant association was found with SNP rs7161123 in the fucosyltransferase 8 (FUT8) gene ($p=1.09x10-8$), while SNP rs3020450 in estrogen receptor beta (ESR2) also showed significant association ($p=6.56x10-6$). Interestingly, stratification to sex revealed that FUT8 is restricted to females, while ESR2 affects the glycan in both sexes. Moreover, FUT8 seems to be associated only in pre-menopausal females, while ESR2 is more associated in post-menopausal women [23]. This is the first study in which human glycan patterns could be associated with SNPs in glycosylation-related genes, however, replication

of this study is necessary given the relatively low sample size for a genetic study.
In the last study [24], N-glycans were analyzed from plasma from 1914 individuals (742 males, 1172 females, between 18 and 98 years of age) using HPLC, resulting in 33 glycan features, similar to [21]. Changes in levels of glycan features were observed with increasing age, and were sex specific. In general, females show more profound associations between glycosylation and age than males, while some glycans may even show opposite associations in males compared to females. Interestingly, glycosylation patterns of women changed most dramatically between the age-groups 40-49 and 50-59, implying an influence of the hormonal changes associated with entrance of the menopause. Several glycan features associated with chronological age could also be associated with body fat parameters, however, in a multivariate analysis including both groups of parameters, it was revealed that mainly tetra- and trisialylated compounds correlate positively with cholesterol and lipoproteins. Changes in glycosylation could also be observed in smoking individuals. Interestingly, only small amounts of the glycosylation variance could be explained by the monitored calendar age and few environment related parameters.

Few studies on altered glycosylation associated with diabetes mellitus have been seized. In a recent study, plasma glycosylation profiles of twenty Type 2 diabetes patients and eighteen healthy controls were compared [25]. Glycans were released, labeled using PA and subsequently analyzed using RP-HPLC with fluorescence detection. It could be reported that levels of glycans carrying α1-6 linked fucose are increased in patients with Type 2 Diabetes Mellitus [25].

Several groups have reported on the changes in protein glycosylation with specific types of cancer (e.g. [26-38]). In a study on altered glycosylation with breast cancer, glycans from serum proteins from 27 healthy individuals and 82 patients with breast cancer were released and permethylated prior to analysis by MALDI-MS [38]. Increased levels of sialylated as well as fucosylated glycans were observed in the cancer patients. Altered glycosylation patterns associated with overian cancer have also been studied in a larger scale study [33]. Glycans from serum samples from 90 con-

trols as well as 90 diseased individuals were analysed using HPLC, and increased levels of the core-fucosylated, non-galactosylated biantennary glycan, as well as glycans carrying antenna fucosylation on sialylated antennae were reported.

Glycan-based biomarkers in specific glycoproteins

Besides changes in total body fluids, it is possible to analyse the glycosylation patterns of specific glycoproteins from these body fluids, if it is expected that such proteins are more specific markers for a certain physiological state. Several papers on specific protein glycosylation have been published, with IgG glycosylation being mostly studied. We recently published an overview of disease states in which plasma IgG glycosylation is altered [39].

Several papers have described the changes in immunoglobulin G glycosylation associated with chronological age. In 1988, the first study on IgG glycosylation with chronological age was published. The levels of galactosylation on IgG were monitored in plasma of a group of 151 normal, healthy individuals of both sexes varying in age from 1 to 70 years. An increase in galactosylation was observed until the age of 25, while the galactosylation was found to decrease at older age [40].

Approximately ten years later two IgG glycosylation studies in healthy individuals were published by research teams from Japan. Yamada et al. enzymatically released and analyzed N-glycans from human serum IgG of 176 female and 227 male individuals ranging in age from 0 to 84 years [41]. The age-dependency of galactosylation could be confirmed for both sexes. In addition, a difference in the degree of galactosylation was found between males and females at an age of approximately 25 years: women showed on average lower levels of agalactosyl glycoforms than men. Moreover, the occurrence of bisecting N-acetylglucosamine showed an age-dependency: bisecting GlcNAc is generally increasing with age and seems to reach a plateau at 50 years of age.

Corroborating the findings of Yamada et al. [41], Shikita et al. (1998) released N-gly-

cans from human serum IgG of a small cohort consisting of 43 female and 37 male individuals ranging in age from 18 to 73 years [42]. The age-dependent decrease in galactosylation could only be confirmed for females, while for males only a slight trend towards lower galactosylation at higher age was observed. However, due to the small number of male individuals in the study, this tendency was not found to be significant. The incidence of bisecting GlcNAc was found to increase with age for both sexes [42].

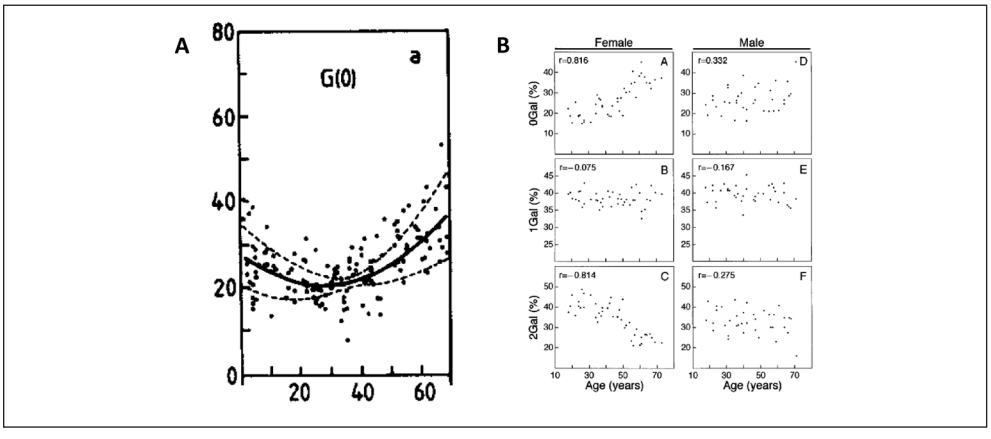

Figure 1-3. *Association of IgG glycosylation with calendar age. A. level of galactosylation first increases with calendar age, and decreases after approximately 25 years of age, reprinted with permission from [40]. B. Association of galactosylation with calendar age is sex specific, reprinted with permission from [42].*

A decreased IgG galactosylation was not only associated with higher calendar age, but also with conditions of poor health. Decreased levels of galactosylated glycans on immunoglobulin G in patients with the autoimmune disease Rheumatoid Arthritis (RA) were reported for the first time in 1985 [43]. Immunoglobulin G was captured from serum from 18 patients and 15 control individuals and glycans were analyzed after hydrazynolysis and decreased levels of galactosylation on the α1-6 antenna. Similar results could later be obtained in other studies [44-46], as well as for several other inflammatory diseases (e.g. [47]).

Altered glycosylation patterns of the specific plasma protein α-1 acid glycoprotein (AGP) have been associated with diabetes mellitus. AGP is an acute phase protein produced in the liver and is heavily glycosylated. In a first study, comprising 39 Type 1 diabetic patients and 24 non-diabetic individuals as controls, the glycosylation of AGP was studied [48]. Using Aleuria Aurantia lectin (AAL), which binds fucosylated glycans, in crossed affinity iimunoelectrophoresis, in combination with high-pH anion exchange, it was observed that there is an increase in antenna fucosylation of AGP in plasma from patients with Type 1 Diabetes Mellitus [48]. Later Higai et al. [49] studied the glycosylation of serum derived AGP in five Type 2 Diabetes patients compared to five healthy controls by glycan release, permethylation and subsequent MALDI-TOF-MS analysis. They observed a tendency towards increased fucosylation in Diabetes patients, but no significant results were obtained.

In the search for biomarkers that reflect Alzheimer's disease |(AD) in the easily available body fluid plasma, the glycosylation patterns of both plasma- as well as CSF-derived reelin, a large glycoprotein, which has been associated with neurodegenerarion, were investigated [18]. The glycosylation pattern of Reelin from both CSF and plasma from 19 patients with AD and 11 nondemented controls was examined using binding assays with several lectins. Decreased binding of CSF derived reelin to the Lens culimaris agglutinin-A (LCA) lectin (which binds core fucose) with AD was observed, however, this could not be corroborated in plasma.

In this thesis a total plasma screening (chapter 5) was followed by the analysis of the glycosylation patterns of Immunoglobulin G (chapter 6) and Immunoglobulin A and alpha-1-antitrypsin (chapter 7). While Immunoglobulin G and Immunoglobulin A are representatives of glycoproteins produced in B-lympocytes, alpha-1-antitrypsin is produced in the liver by hepathocytes. It is therefore expected that the immunoglobulin glycosylation is differently regulated compared to the glycosylation of liver derived glycoproteins.
IgGs are present in high abundance in the human circulation (concentration approximately 10 mg/ml in human plasma) and occur in four different subclasses (IgG1,

IgG2, IgG3, and IgG4 in humans). Since more than two decades, human IgGs are known to vary in the N-glycosylation at the highly conserved N-glycosylation site of the Fc part (e.g. [43]). Human IgAs exists in the two subclasses IgA1 and IgA2. Plasma IgA consists for 90% of IgA1, and 10% IgA2, while secretory IgA may consist of up to 50% IgA2 [50]. IgA1 carries two N-glycosylation sites, while up to 5 N-glycans may be attached to IgA2. IgA1 moreover, possesses a hinge region carrying up to 5 O-glycans. Secretory IgA contains secretory component, which carries an additional seven N-glycosylation sites, as well as J chain with one N-glycosylation site [50]. In contrast to the immunoglobulins, alpha-1-antitrypsin is mainly produced in the hepatocytes, culminating in a plasma concentration of 1.0-2.8 mg/ml. The protein carries 3 N-glycosylation sites, which mainly carry bi- and tri- antennary glycans [51].

Glycosylation analysis

The analysis of the glycan moiety attached to proteins can be performed at three different levels (see Figure 1 4): The intact glycoprotein, glycopeptides generated using different proteinases, or released glycans can be studied. In this thesis two of these approaches were applied. First of all, glycopeptide analysis was performed to profile glycosylation features of IgG, as IgG glycopeptide analysis is a widely established technique, which has, also in our lab, been used quite extensively (e.g. [52;53]). As analysis of released glycans is more widely applicable, this approach was optimized in this thesis for high-throughput analysis, and subsequently applied for most of the studies in this thesis.

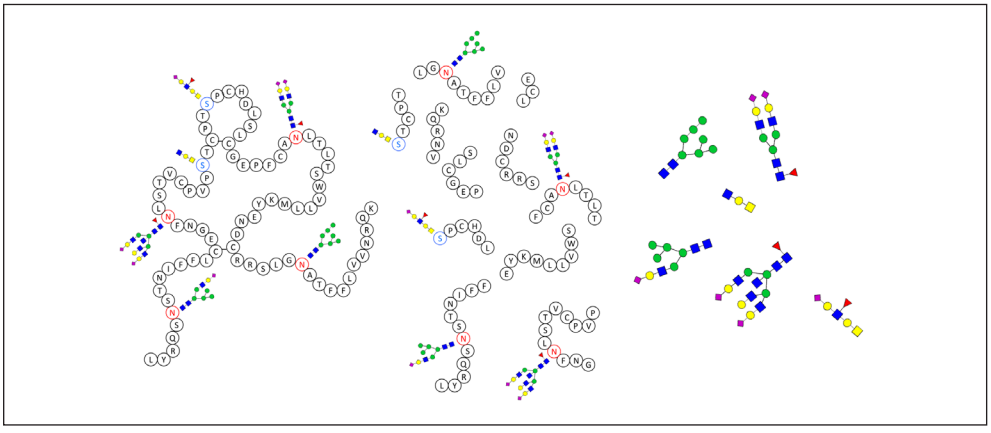

Figure 1 4. *N-glycan analysis can be performed at three different levels. The intact protein (left), the (glyco) peptides (middle), or the released glycans (right) may be studied.*

Analysis of intact glycoproteins

A major advantage of the analysis of intact glycoproteins is the minimal need for sample preparation. After purification, the intact glycoprotein is directly accessible for analysis. Moreover, mass spectrometry of intact glycoproteins provides an integrated analysis of a large number of PTMs. Besides glycosylation these include e.g. phosphorylation, the conversion of glutamine into pyroglutamate, oxidation on methionine residues, deamidation and formation of succinimide from aspartic acids, cysteinylation, dimer formation, and peptide bond cleavage, e.g. by acid hydrolysis, which can also be induced by analytical conditions (e.g. acidic conditions at elevated temperature in LC [54]). This approach is, however, not feasible for more complex mixtures of glycoproteins because of the enormous heterogeneity, which prohibits mass spectrometric detection [39].

IgG glycopeptide analysis

Mass spectrometric analysis of glycoproteins - and particularly the IgG subclasses - at the level of glycopeptides which are generated with endoproteinases offers the possibility to determine glycan attachment sites and evaluate glycosylation site

micro-heterogeneity. Usually these analyses are performed on purified proteins [55-58]. By far the most popular technique for IgG purification is affinity-purification using immobilized Protein A (e.g. [54;59]) and Protein G (e.g. [41;60-62]): these bacterial proteins bind to IgGs of various species and are commercially available in immobilized form (e.g. Sepharose beads). For the purification of IgG from human serum, both Protein A and Protein G may be applied, with the important difference that Protein G will bind all 4 human IgG subclasses, while Protein A will only retain IgG1, IgG2, and IgG4, but not IgG3 [52;63]. Various specific and nonspecific proteolytic enzymes are available for the digestion of glycoproteins, such as trypsin, lysylendopeptidase (Lys-C), endoproteinase Glu-C (V8 protease) and pronase; their activity is usually enhanced by protein denaturation.

Trypsin is by far the most widely applied proteolytic enzyme in proteomics and glycoproteomics. It is a serine protease, cleaving peptides at the carboxyl side of lysine and arginine, unless the following amino acid is a proline. IgG Fc glycopeptides resulting from tryptic digestion are not extremely basic or acidic and are often easy to ionize for mass spectrometric detection. The analysis strategies for the generation of profiles from such glycopeptides are thus relatively facile and straight forward. Therefore, the IgG-Fc glycosylation profiles generated in the studies in this thesis were obtained using tryptic glycopeptides. Due to the occurrence of tryptic Fc glycopeptides with identical peptide moieties for the IgG2 and IgG3 subclasses, the analysis of a proteolytic digest of total plasma IgG will not allow the specific assignment of glycopeptides to these subclasses [52]. To overcome this problem, IgG1, IgG2, and IgG4 may first be captured using Protein A, while IgG3 can subsequently be isolated from the flow-through using Protein G.

To allow determination of IgG-Fc glycosylation profiles in a site-specific manner, mass spectrometric detection is necessary. Nowadays, electrospray ionization (ESI)-MS or matrix assisted laser desorption (MALDI)-MS are the method of choice. However, both techniques are highly susceptible for impurities (especially salts) in the sample and therefore proper sample pretreatment is important. Sample purifica-

tion may be achieved by the application of hyphenated techniques such as LC-ESI-MS [52], but samples may also be fractionated and desalted by SPE prior to mass spectrometry [58].

One of the most widely applied analytical techniques for glycopeptide analysis involves direct analysis of proteolytic digests by RP-HPLC coupled to ESI-MS. RP-HPLC separation of proteolytic IgG digests is mostly performed on C18 analytical columns at the micro- or capillary-scale, while nano-LC is, due to impaired method stability less frequently used [39;64]. Alternatively, graphitized carbon chromatography coupled to MS detection [65] or Zwitterionic interaction chromatography - hydrophilic interaction chromatography (ZIC-HILIC)-LC-MS [66] may be applied. ESI-MS/MS spectra of LC separated glycopeptides can be identified on the basis of glycan specific oxonium ions of m/z 204 [HexNAc+H]+, 366 [$Hex_1HexNAc_1$+H]+, 528 [$Hex_2HexNAc_1$+H]+, and 657 [$Hex_1HexNAc_1NANA_1$+H]+ obtained on collision induced dissociation (CID) [39;67-69]. Moreover, multiply charged IgG glycopeptide ions as generated by ESI allow the performance of electron transfer dissociation (ETD) experiments, resulting in peptide sequence information [52].

IgG-Fc glycopeptides are often studied by MALDI-MS. Tryptic glycopeptides are overlaid, mixed or sandwiched with 2,5-dihydroxybenzoic acid (DHB) and/or α-cyano-4-hydroxycinnamic acid (CHCA) and analysis can be performed by MALDI-Time Of Flight (TOF)-MS in linear and/or reflectron mode. Due to the soft nature of MALDI ionization, most IgG glycopeptide ions remain intact. However, sialylated glycopeptides often show massive desialylation due to in-source decay and metastable decay when analyzed by positive-ion MALDI-TOF-MS with delayed extraction in the reflectron mode. In contrast, detection of sialylated glycopeptides can be achieved by MALDI-TOF-MS in the linear mode [53;55], but this results in poor resolution.

Glycan analysis – sample preparation

An efficient, reliable method to release the oligosaccharides from their protein or peptide backbones is a prerequisite for the detection of the glycans [70]. Several ap-

proaches for N-glycan release have been reported, which are represented in Figure 1 5.

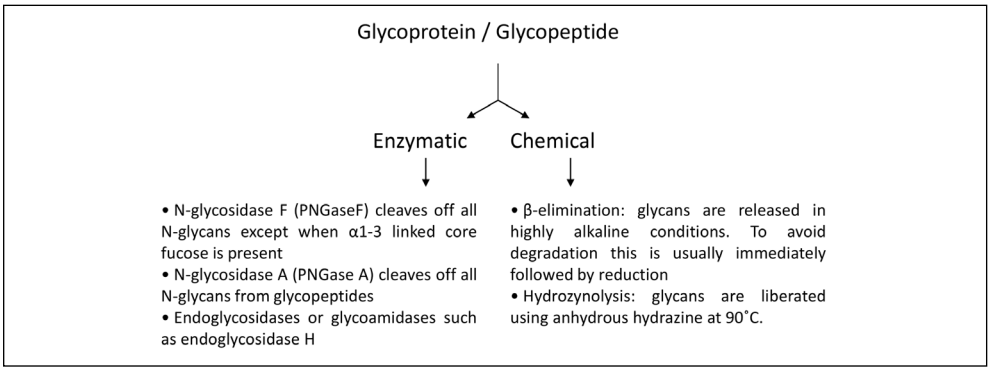

Figure 1-5. *Strategies for N-glycan release from glycoproteins and glycopeptides. Glycan may be released either enzymatically or chemically. The Figure is modified from [71] and reprinted with permission.*

N-glycans are commonly released enzymatically using PNGase F or A [70] (see Figure 1 5) as this is the easiest and most reliable approach, resulting in glycans with a free reducing end. However, other release methods have also been employed. Several protocols for the use of β-elimination for the release of O-glycans have been developed (e.g. [26;72-75]), but all protocols also release N-glycans to a substantial degree. β-Elimination could thus be envisioned to be a perfect method for combined N- and O- glycan analysis; however, a major drawback is that the reaction is hampered by water and usually results in reduced-end glycans, which are hard to derivatize for analysis. Attempts have been made to develop a non-reductive β-elimination [72;74], however, such methods have not been widely applied in glycan analysis. The hydrazinolysis approach is also applicable for the release of N- as well as O-glycans [70;76;77] yielding glycans with a free reducing terminus; however, the use of anhydrous hydrazine, a highly toxic and explosive chemical also used in rocket fuel [78], hampers the application of this approach, especially at a high-throughput level.

Reductive amination

Upon release, the N-glycans may be derivatized to facilitate their analysis. Several

derivatization approaches have been described, among which reductive amination Figure 1 6 (e.g. [20;21;42;79-81]), permethylation [82-84], Michael addition [85] and hydrazide labeling [86] are most widely applied. In the studies in this thesis, only reductive amination has been employed, as most other labeling techniques are less suitable for high-throughput analysis. In reductive amination, glycans are labeled at their reducing end: a label containing a primary amine group reacts in a condensation reaction with the aldehyde group of the glycan, resulting in an imine or Schiff base, which is reduced by a reducing agent to yield a secondary ami

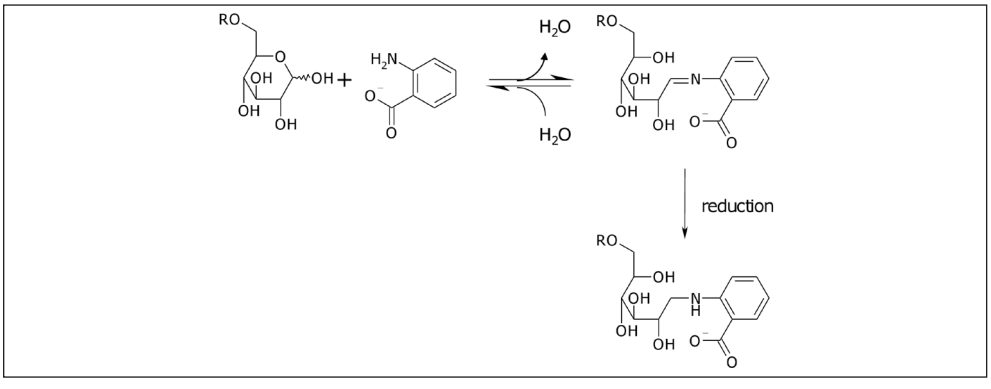

Figure 1-6. *Reaction scheme of derivatization of N-glycans by reductive amination. In the scheme 2-aminobenzcacid is used as the label.*

The reaction is often performed in dimethylsulfoxide (DMSO) containing acetic acid [87], but alternative approaches using tetrahydrofuran [88] and methanol [89] have been described. An advantage of this labeling approach is the stoichiometric attachment of one label per glycan allowing a direct quantitation based on fluorescence or UV-absorbance intensity.

Labels

Various labels have been used for the reductive amination of glycans. In two reviews, an extensive, though not complete list of labels for reductive amination has been given [90;91]. The most widely applied labels are 2-aminobenzamide (2-AB),

2-aminobenzoic acid (2-AA), 2-aminopyridine (PA), 2-aminonaphthalene trisulfonic acid (ANTS), and 1-aminopyrene-3,6,8-trisulfonic acid (APTS), which are depicted in Figure 1 7.

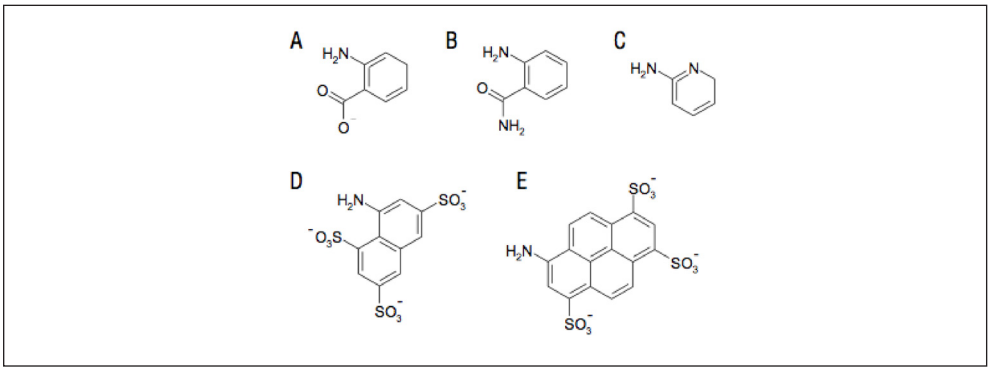

Figure 1-7. *Structures of the most widely applied labels for reductive amination of oligosaccharides. 2-AB (A), 2-AA (B), PA (C), ANTS (D) and APTS (E).*

Labeling kits are available for the tags 2-AB, 2-AA and PA (see e.g. www.ludger.com) as well as for labeling with APTS (see www.beckmancoulter.com) and ANTS (see www.prozyme.com). 2-AB is a label that lacks negative charges and is widely applied in chromatographic analysis. An extensive database has been developed which uses the standardized elution positions of 2-AB labeled glycans in HILIC with fluorescence detection for structural assignment [80]. PA is also widely used in HPLC profiling [66;92], and databases for structural assignments based on standardized elution positions have been developed as well [93]. A major drawback of the use of PA is the necessary recrystallization step prior to use when the commercial labeling kit is not used, as the commercially available crude compound lacks sufficient purity. The 2-AA label carries one negative charge, which makes it very versatile. It is used in HPLC and CE separations as well as in positive-mode and negative-mode MALDI analysis allowing detection of both neutral and sialylated glycan species [94-97]. APTS has three negative charges and is therefore very suitable for capillary electrophoresis (CE) and capillary gel electrophoresis (CGE; see below) [98;99]. However, the analysis of APTS-labeled oligosaccharides by MALDI appears to be difficult [100]. In the studies presented in this thesis 2-AA (Chapters 2, 3 and 5), 2-AB

(Chapter 3) and APTS (Chapters 4 and 7) have been applied.

Reducing agent

The most widely applied reducing agent is sodium cyanoborohydride. Very high yields of labeled oligosaccharides have been reported using this compound [87]. Alternatively, sodium triacetoxyborohydride [101] and borane-diethylamine [102] have been used; however, these compounds are not widely used within the glycoanalytical field.

As an alternative, the use of the non-toxic chemical compound 2-picoline-borane for the reductive amination of oligosaccharides with 2-AA and 2-AB is described in this thesis (Chapter 3, [103]).

Optimal reaction conditions

The optimal conditions for the reductive amination reaction of N-glycans with the 2-AB and 2-AA label using $NaBH_3CN$ as the reducing agent were reported by Bigge et al. in 1995 [87]. The concentration of the labeling agent was recommended to be ≥ 0.25 mol/L, while more than 1 mol/L reducing agent was required. The derivatization is enhanced by the addition of glacial acetic acid up to a content of 30% (v/v). A reaction temperature of 60˚C was found to be optimal. Using these conditions, all glycans are derivatized within 2 hours, and glycan degradation reactions such as acid-catalyzed loss of sialic acid are minimized. In a different study, the influence of different acids was evaluated for the derivatization of oligosaccharides with APTS [104]. Acids with lower pKa values, like citric acid, malic acid, and malonic acid were shown to result in higher yields than acetic acid when the derivatization was performed at 37˚C for 16 h. When higher temperatures and very acidic conditions are applied, partial degradation of the glycans will be induced with loss of sialic acids [105]
Strategies for purification

After derivatization, glycans often have to be purified prior to analysis. Not only the excess of salts has to be removed (e.g. for MALDI analysis), but also the concen-

tration of the labeling reagent, normally present in large excess during the labeling step, should be reduced. Five major strategies are applied for this purpose: solid phase extraction (SPE) [75;80;97;106;107], liquid-liquid extraction [108], gel filtration [46;109], paper chromatography [110] and precipitation [107]. Typical examples of purification strategies are summarized in Table 1 1. SPE is mainly performed after reductive amination, while liquid-liquid extraction is the method of choice after permethylation [82;108]. In permethylation procedures, liquid-liquid extraction may be followed by C18 SPE [111]. Paper chromatography may be used after reductive amination and hydrazide labeling [77;112]. Precipitation with acetone is not widely applied, but may be used to remove proteins [107]. The strategies strongly differ in complexity as well as time consumption of sample handling. Among the five strategies, SPE is the method of choice for analyzing larger sample cohorts as it can be adapted to high throughput set-ups [97]. The diversity of stationary phases for SPE is relatively large, and its use in purification of derivatized glycans is extensive. Several stationary phases have been used for purification of derivatized glycans by SPE: reverse phase (RP), porous graphitized carbon (PGC), hydrophilic interaction liquid chromatography (HILIC) and anion exchange chromatography (AEX).

Method	Material	Label	Comments	Ref.
SPE	RP, C18	RedAm, aniline		[113]
SPE	RP, C18	Permethylation	TFA may be added	[75;114]
SPE	RP, C18	HL, BACH		[115]
SPE	PGC	RedAm, 2-AB	Sample application in water, elution using 25% (v/v) ACN	[116]
SPE	PGC	RedAm, benzylamine	Sample application in water, elution using 20% (v/v) ACN	[117]
SPE	HILIC, nylon filter, Oasis HLB or amide-2	RedAm, 2-AA	Sample application in 95% acetonitrile, elution using water	[94;118]
SPE	HILIC, cellulose	RedAm, 2-AA	Sample application in 80% acetonitrile, elution using water	[97]
SPE	HILIC, microelution plate	RedAm, 2-AB	Sample application in 80% acetonitrile, elution using water	[80]
SPE	HILIC, DPA-6S	RedAm, APTS	Sample application in 90% acetonitrile, elution using water	[119]
Gel filtration	Sephadex G10	RedAm, APTS	Gelfiltration at 96-well format in filter plates	[99]
Gel filtration	Toyopearl HW-40F	RedAm, APTS		[120]
Gel filtration	Sephadex G15	RedAm, PA		[121;122]
Liquid-liquid extraction	Water: chloroform	Permethylation	Permethylated glycans are in the organic (lower) phase	[82;108; 123;124]
Liquid-liquid extraction	Water: dichloromethane	Permethylation	Permethylated glycans are in the organic (lower) phase	[125]
Liquid-liquid extraction	Water: ethylacetate	HL, phenylhydrazine	Labeled glycans are in the aqueous phase	[86]
Liquid-liquid extraction	Water: chloroform	MA, PMP	Labeled glycans are in the aqueous phase	[126]
Paper chromatography	3 MM Whatman filter paper	RedAm, 2-AB, DAP	Glycans are eluted using water	[77;112]
Precipitation	Acetone	RedAm, PA	Precipitation with water free acetone was conducted three times	[107]

Table 1-1. *Sample purification after glycan labeling. RedAm = reductive amination, HL = hydrazide labeling, MA = Michael addition.*

The introduction of hydrophobic properties by derivatization of native glycans (e.g. using the 2-AB label) is necessary for their adsorption to reverse phase material, since non-derivatized glycans do not show sufficient retention on this material.

PGC is most widely and originally applied for non-derivatized glycans [127], but its use with labeled glycans has also been reported [116;117]. Even though the mechanism is not well understood [128], the optimized method is highly specific for glycans. Elution is usually performed with acetonitrile:water mixtures containing TFA (Table 1 1). Drawbacks of this method are the relatively high costs of the PGC, and the fact that excess label may not be removed, depending on the properties of the label.

An upcoming approach for purification of mainly reductively aminated glycans is HILIC [80;94;97]. Several HILIC stationary phases have been described. Using this approach glycans are retained based on their hydrophilic properties, while less hydrophilic excess label may be removed. In this thesis, the application of cellulose for purification of 2-AA-labeled glycans on a high-throughput platform is described (Chapter 2, [97]), as well as the use of Biogel P-10 for the purification of APTS-labeled glycans (Chapter 4). Since elution of the derivatized glycans can be performed with water, samples can be stored immediately without the risk of the acid-catalyzed hydrolysis of sialic acids.

Glycan purification strategies may also include fractionation by anion exchange chromatography of oligosaccharides labeled with 2-PA or 2-AB allowing the separation of neutral from sialylated glycans [129;130].

Glycan analysis – separation and detection

Glycan analysis is nowadays typically performed using mass spectrometric, fluorescence or pulsed amperometric detection. While fluorescence and pulsed amperometric detection have to be preceded by glycan separation using chromatographic or electromigrative techniques, mass spectrometric detection can be performed in

a stand-alone manner, but may also be hyphenated with an additional separation technique. In this thesis, we describe the use of HPLC-FL, MALDI-TOF-MS, CE-ESI-MS, HPLC-ESI-MS, CGE-LIF and MALDI-FTICR-MS. Several extensive reviews regarding separation and detection of glycans have been published [70;71;131-133] and to give a full overview of all possibilities here would go beyond to scope of this thesis. Therefore, this part of the introduction is mainly focused on the techniques used in this thesis. An overview of the separation techniques with some of their advantages and disadvantages is depicted in Figure 1 8. High-pH-anion-exchange chromatography has been added as this technique is currently widely applied in industrial applications.

HILIC-LC	RP-LC	PGC-LC	HPAEC	CE and CGE
Both native and derivatized glycans	Derivatization necessary for retention	Both native and derivatized glycans	Both native and derivatized glycans, though native glycans are preferred	Derivatization with (preferably negatively) charged tag necessary to induce migration of neutral glycans
Injection in high organics, elution using increasing water content	Injection in water, elution using increase in organic solvent	Injection in water, elution using increase in organic solvent	Injection in water, elution using increase in ionic strength	Injection in water, migration using constant buffer and external voltage
Retention based on glycan, separation based on hydrophilicity	Retention largely dependent on label, separation based on hydrophobicity	Retention based on glycan and label, separation not well understood	Retention based on glycan (and charged label), separation based on charge and size	Migration based on charge and size
Separation in 30 min., no multiplexing	Separation in 30 min., no multiplexing	Separation in 30 min., no multiplexing	Separation in 15 min. possible, no multiplexing	Separation in 3 min. possible, multiplexed system available
Easy to couple to MS	Easy to couple to MS	Easy to couple to MS	Possible to couple to MS using a desalting device	CE possible to couple to MS, however, dependent on buffer system
Not available in most labs	Available in most labs	Not available in most labs	Not available in most labs	Not available in most labs

***Figure 1-8**. Separation techniques most widely applied in glycan analysis.*

Chromatographic separations

HPLC is a widely applied separation technique in glycan analysis. Several stationary phases, including hydrophilic interaction chromatography (HILIC), reverse phase (RP) and porous graphitic carbon (PGC) have been applied. A major disadvantage of such separation techniques is that run-times of at least 30 min. are needed, and multiplexing is not yet available. Therefore, this technique is not very suitable for high-throughput analysis.

HPLC-mode HILIC is a very common technique for oligosaccharide separation [80;97;134]. Hydrogen bonding is often the major retention mechanism. In order to be effective in a pure HILIC mode (without anion exchange present), solvents with a rather high ionic strength have to be used, resulting in the suppression of ionic interactions of acidic groups on the glycan with the ammonium groups of the stationary phase. HILIC is sometimes referred to as size separation, due to the correlation between glycan size and its retention time. However, the contribution of added monosaccharides to oligosaccharide retention may vary considerably, depending on the composition of sugars of the glycan. Generally speaking, hexoses tend to give a larger (increase in) retention than N-acetylhexosamines or fucoses [80;110]. Rudd and coworkers have established a database (GlycoBase and autoGU) which allows the structural assignment of 2-AB-labeled oligosaccharides based on migration positions on a TSK-amide80 column. In this database, retention times normalized to an axis of glucose units (GU) are listed for 2-AB-glycans before and after cleavage by exoglycosidases [80;135]. HILIC also has a considerable potential to separate structural isomers [136-138].

RP-HPLC of oligosaccharides requires glycan labeling, as the retention of underivatized glycans is often insufficient [139]. For labeled glycans, the retention properties are determined by a combination of the hydrophobic properties of the tag and the contribution of monosaccharide residues [107]. The effect of the addition of monosaccharides on the retention of a labeled glycan may differ depending on the tag: N-glycans labeled with pyridylamine (PA), which is a low-retention tag, showed increa-

sing retention with a higher degree of antenna galactosylation, whilst the same set of N-glycans labeled with ABEE (2-aminobenzoic acid ethyl ester) showed decreasing retention with a higher degree of antenna galactosylation [107]. The separation power of RP-HPLC for 2-AB-labeled glycans was recently demonstrated by Chen and Flynn [140], who could separate several structural isomers.
Porous Graphitic Carbon (PGC) stationary phases are often used for the separation of oligosaccharide alditols [141], fluorescently labeled glycans, and even permethylated glycans [83], as reviewed recently [128]. Retention is not well understood, but is mediated by a combination of predominantly ionic and hydrophobic interactions [128]. Notably, PGC is very efficient in separating structural isomers. Retention times in combination with glycan mass analysis are often sufficient for structural assignment [64;142].

A widely applied alternative for HPLC separation is high-pH anion exchange chromatography (HPAEC), which is performed at a pH of approximately 13. At this pH, hydroxyl groups on glycans are partially deprotonated, resulting in a partial negative charge, which is used for the separation of the glycans. HPAEC is suitable for the separation of non-derivatized glycans as well as glycans with various types of labels and modifications. Glycans with a reduced end exhibit less retention than reducing end-glycans, as they lack the anomeric hydroxyl group which is a major contributor to the partial negative charge [71]. Usually pulsed amperometric detection is applied for the detection of glycans, however, with the aid of a desalting system, hyphenation with MS is possible and will be more widely applied [143].

CE and CGE

Electromigrative separations (e.g. CE and CGE) are based on differences in the effective electrophoretic mobility of the analytes in the separation medium, when applying an external voltage. While “acidic” glycans carrying e.g. sialic acid are negatively charged already at low pH (pKa of sialic acids 2.0-2.8 [144]), “neutral” glycans will only exhibit partial negative charges at very high pH due to the deprotonation of hydroxyl groups. Therefore, the introduction of a (negative) charge by derivatization

is necessary to separate all N-glycans in one run. CE gives rise to a charge-based separation, therefore sialoforms are relatively easy to separate with high resolution [97]. An interesting approach for the separation of APTS-labeled glycans was presented by Zhuang et al. [145]. The authors used chip technology to allow a very fast analysis (< 3 min) with high resolution.

A number of methods for the separation of 2-AA and APTS-labeled glycans using CGE have been published, mainly including polyethylene glycol as sieving matrix [96;146;147]. Multiplexing using a DNA sequencer has been developed for high-throughput glycoprofiling of APTS-labeled glycans [98;99;120;148;149], and therefore this technique is very suitable for high-throughput profiling. The major disadvantage of CGE-based methods is their incompatibility with MS, so that the identification of signals is complicated and can only be achieved by standards or additional experiments such as the re-analysis of the sample after exoglycosidase treatment [96;98].

Detection

Detection of glycans after separation by HPLC and CE may be performed using fluorescence detection, pulsed amperometric detection or mass spectrometry. Detection using fluorescence requires the introduction of a fluorophore by derivatization, however, this immediately facilitates quantitation, as one fluorophore is added to each glycan. Laser induced fluorescence detection is necessary after CE or CGE separation, as the optical window of such separations is very small (e.g. [96;99]. A major disadvantage of non-MS-based detection techniques is the lack of direct possibilities for identification. However, when profiling large datasets of similar samples, the amount of data is reduced using non-MS based detection and identification is of less importance as similar glycans will appear in each profile. MS is usually performed using electrospray ionization (ESI) (e.g. [116;141]). Fractionation of glycans for structural analysis is may then be performed using collision induced fragmentation (CID).

Direct MALDI-MS

Direct mass spectrometry may also be applied for the analysis of glycans. Even

though direct infusion ESI-MS could be applied, Matrix Assisted Laser Desorption ionization (MALDI)-MS is much more widely applied and 2,5-dihydroxybenzoic acid (DHB) is almost exclusively applied as the matrix for MALDI ionization of oligosaccharides. Ionization efficacies of oligosaccharides are markedly influenced by charged substituents of the glycan such as carboxyl groups of sialic acid and derivatization may also influence ionization efficacies by introducing additional charges. Glycans with negatively charged substituents are more efficiently ionized in negative-ion mode than their neutral counterparts, whilst native as well as labeled glycans lacking acidic substituents are predominantly analyzed using positive-mode ionization [150]. Introduction of a negative charge (e.g. by derivatization using 2-AA), makes neutral glycan structures readily ionizable, thereby allowing the joint registration of sialylated and non-sialylated glycans in one MALDI spectrum, as demonstrated in Chapter 2.

When analyzing sialylated glycans, reflectron mode MALDI-TOF-MS often shows partial or complete loss of sialic acids which is due to in-source decay and metastable decay. MALDI-TOF-MS instruments generally require source pressure in the range of 10-6 mbar. The use of instruments with intermediate pressures (typically sub mbar range) suppresses the loss of sialic acids, as loss of sialic acids is avoided by collisional cooling of ions in the ion source at intermediate pressure conditions. As the analysis time of one sample by MALDI-MS is generally below 1 min., and relative quantitation is possible, this analysis strategy is well-suited for high-throughput analysis.

Human longevity

All human individuals age; however, some individuals age faster than others, and some individuals show extreme longevity [151]. In animal models lifespan extension could be initiated by applying calorie restriction and by genetic mutations in single jeans. Lifespan extension in animal models is driven by alterations in nutrient sensing pathways such as Insulin-IGF signaling (IIS) and TOR signaling. Bartke [152;153] recently reviewed the association of the insulin/IGF pathway with longevity in several species. As it is difficult to separate the roles of insulin and IGF-1 in mammals, con-

troversy has risen regarding their individual contribution to longevity. Interestingly, insulin sensitivity was increased upon calorie restriction in humans, a state which is generally regarded as life prolonging [154]. Secondly, it has been suggested that the hypothalamo-pituitary-thyroid axis modulates the process of aging. Low levels of thyroid hormone have been associated with delayed aging in mice. More evidence was found in humans: in centenarians high levels of thyrotropin were associated with prolonged life span [155]. A third pathway, which has been associated with longevity, is cellular repair mechanisms and cellular responses to stress. So far, evidence has come from model organisms: between species, fibroblasts from longer living animals have a higher resistance to stress in vitro [156;157] and fibroblasts from long-lived mice were more stress resistant than their wild-type counterparts [158].

The Leiden Longevity Study

All samples studied in this thesis originate from the Leiden Longevity Study (LLS). The Leiden Longevity study is a cohort based on families of exceptional longevity and consists of 420 Caucasian families with at least two long-lived siblings. The study is illustrated in Figure 1 9. Long-lived is in this study defined as over 89 years of age for males and over 91 years of age for females. In the Netherlands in 2002, the percentage of male inhabitants over 89 years was 0.5% and the percentage of female inhabitants over 91 years was also only 0.5% [159]. As the Leiden Longevity study aims to address familial longevity, only siblings which were long-lived were included and it is estimated that far less than 0.1% of the Dutch population fulfills these criteria [159]. Next to the long-lived siblings (n = 960), also their offspring (n = 1710) and the partners of the offspring (n = 761) were recruited, to allow case-control studies.

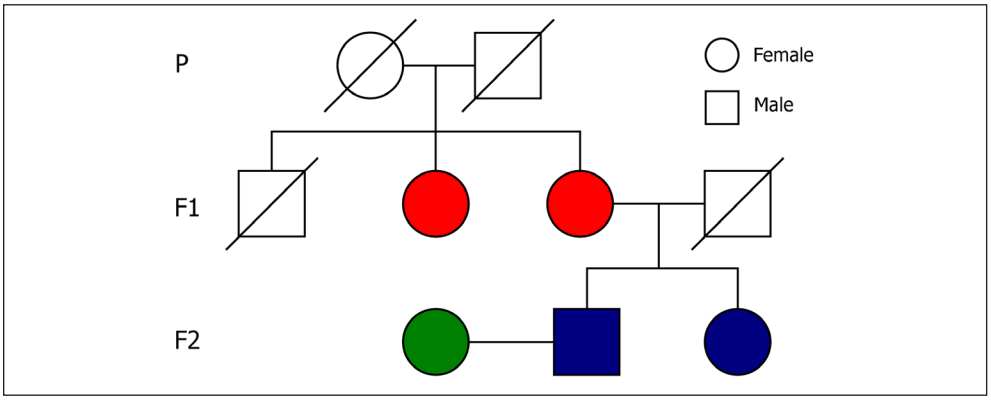

Figure 1-9. *Design of the Leiden Longevity Study. Longlived siblings are depicted in red, their offspring in blue and the partners of the offspring in green.*

In a first study on a subpopulation of the LLS, a 30% decreased mortality risk was observed in the survival analysis over three generations (Parents of the siblings (P), their deceased siblings (F1) and the offspring of the siblings (F2)), while mortality in the partners of the long-lived siblings was not significantly affected [160]. These data strongly suggest that the selected population is enriched for familial longevity.

To identify biological parameters that mark familial longevity (biomarkers), the offspring can be compared to their partners. In such a comparison, the offspring is regarded to represent individuals with a more healthy and slowly aging phenotype and a higher susceptibility to become long-lived, while their partners, representing the general population, serve as controls. Both groups have similar ages (59.4 ± 6.5 for offspring, 58.9 ± 7.4 for their partners) and similar adult environmental influences, thus reducing the risk for confounding results.
Recent studies revealed that the prevalence of myocardial infarction, hypertension and diabetes mellitus was significantly lower in the offspring of the nonagenarian siblings [159]. Moreover, non-fasted serum glucose levels were lower in the offspring, after exclusion of all diabetics, which is in accordance with the lower prevalence of diabetes mellitus in this group [161]. Offspring have larger LDL-cholesterol particles than their partners [162], another marker indicating that longevity in these

families associates with beneficial metabolic profile. Interestingly, insulin-like growth factor 1 (IGF-1) and insulin-like growth factor binding protein 3 (IGFBP3) were not altered with familial longevity. Therefore, differences in glucose handling cannot be explained by differences in IGF-1 or IGFBP3 [161], as may have been expected from research in animal models, where IIS signaling drives lifespan extension. The hypothalamo-pituitary-thyroid axis has also been addressed within the LLS: in the offspring lower free thyroxine and triiodothyroxine levels were observed [163]. These findings suggest that the thyroidal sensitivity for thyrotropin is decreased with familial longevity. All of these features resemble the phenotype of calorie restricted animals. This includes the increased resistance to cellular stress, which was also observed in skin fibroblasts for offspring donors in the LLS [164].

Aim of this thesis.

In previous studies altered glycosylation was shown to be associated with many (age-related) major diseases, and the glycan pattern of immunoglobulin G changes with increasing calendar age. We hypothesize that familial longevity is in essence a delayed aging phenotype. Therefore our aim is to search for changes in glycosylation that are associated with familial longevity in the Leiden Longevity Study. As the LLS comprises several thousands of samples, there existed a need for high-throughput analysis techniques to allow large scale glycan analysis. Thus, development of such techniques –especially at the level of sample preparation- which are also applicable for other study cohorts, has been an essential part of the studies described in this thesis.

Strategy and outlook of this thesis.

The research strategy that has been the basis for the thesis is depicted in Figure 1 10.

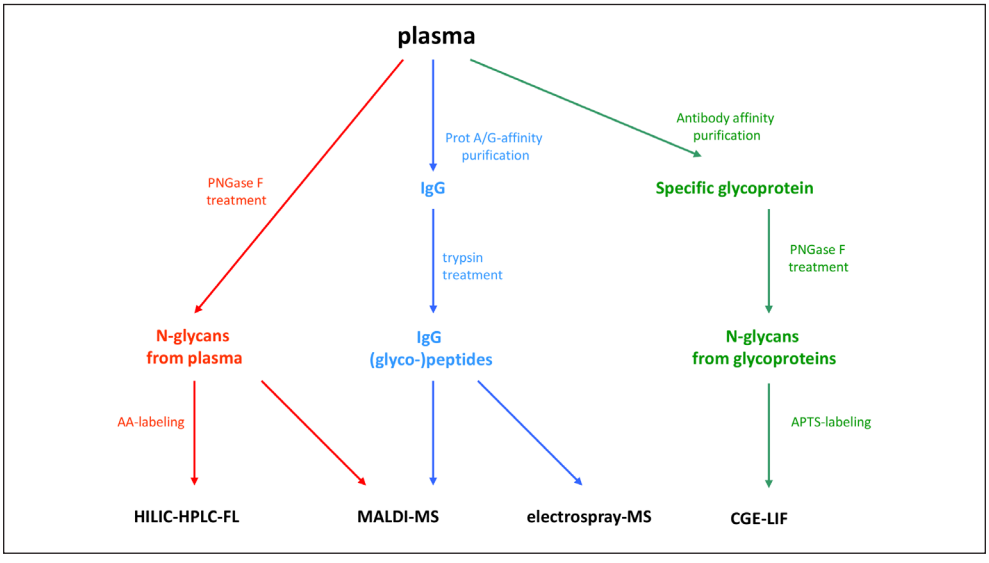

Figure1-10. *Schematic overview of the research strategy used in this thesis.*

In chapter two, a method for sample preparation and N-glycan analysis, based on previous work using HILIC-HPLC [80;165] as well as MALDI-TOF has been developed for high-throughput application (the 'red strategy' in Figure 1 10). In chapter three, the use of a novel, non-toxic chemical compound facilitating glycan derivatization is introduced. In chapter four, the use of a DNA-sequencer for high-throughput analysis of N-glycans is described ('green strategy' in Figure 1 10). Using this analysis method, large amounts of samples could be analyzed in a relatively short period of time. Total N-glycan profiles were obtained from the participants from the LLS using both the HILIC-HPLC method described in chapter two and MALDI-FTICR-MS. Evaluation of these data is described in chapter five, while the profiling and analysis of immunoglobulin G glycosylation patterns is described in chapter six (the 'blue strategy' in Figure 1 10). A method for immunocapturing of α-1-antitrypsin and IgA, in combination with analysis using the DNA-sequencer was described and applied to the LLS in chapter seven. Finally, the results are discussed in chapter eight.