Thesis for the degree of doctor of philosophy (Ph.D.)

# COMPREHENSIVE *N*-GLYCOSYLATION ANALYSIS BY HIGH PERFORMANCE LIQUID SEPARATION TECHNIQUES

by Stefan Mittermayr

Supervisor: Dr. András Guttman


University of Debrecen
Doctoral School of Molecular Medicine

Debrecen, 2013

*to my family*

# TABLE OF CONTENTS

## CHAPTER 4



## CHAPTER 5



## CHAPTER 6



## CHAPTER 7



## CHAPTER 8

# CHAPTER 1

## Introduction and Motivation

Glycosylation is a diverse but critically important post translational modification that modulates the physical, chemical and biological properties of proteins. The discovery of different glycans' involvement in protein confirmation or function, cell-cell signaling and reflection of cellular or even organism physiological state has attracted major research attention. Alterations in glycosylation and the discovery of associated potential oligosaccharide-biomarkers were noted in a number of diseases including cancer.

Due to the analytical complexity associated with glyan analysis, a variety of orthogonal analytical techniques is used, often in combined fashion, and generally including separation driven methods. Separation techniques such as HPLC and capillary electrophoresis (CE) hold an advantage over mass spectrometry based analyses due to their ability to separate isobaric oligosaccharide structures along with more reliable quantitation. Particularly hydrophilic interaction liquid chromatography (HILIC) has consolidated its importance in glycosylation analysis over the past 25 years (initially referred to as normal phase chromatography) due to robustness, reproducibility and the availability of databases that aid elution time based structural annotation. However, a limitation of HILIC based analyses is the long separation time. New sub-2 μm HILIC phases for ultra performance liquid chromatography based glycomics facilitate increased efficiency and significantly decreased analysis time. The ability to separate biologically significant linkage and conformational isomers at higher efficiency is of particular interest in the identification of disease associated glycosylation changes.

CE represents another rapid alternative analytical technique, offering differential electric field mediated separation selectivity. The ability of separating both linkage and positional isomeric species based on inherent molecular shape differences, renders CE as a powerful tool for fine glycan structural analysis. However, robust and reproducible endeavors with associated extensive mapping databases that enable confident structural inference from normalized detection times of fluorescently derivatized glycans yet need to be introduced. The development of a practical analytical method including eligible normalization approaches is crucial to ensure experimental precision as well as structural elucidation accuracy, matching those of the well established HILIC approaches. Finally, the availability of two established orthogonal separation techniques such as HILIC and CE holds the promise to further increase glycan structural elucidation accuracy and confidence.

# CHAPTER 2

# Theoretical Background

## 2.1 Glycosylation represents a complex but very common protein modification

Most natural biopolymers such as oligo-nucleotide or -peptide molecules can be considered as flexible string-like structures due to the simple linear connection of their fundamental building blocks. Monosaccharides, the basic units of oligosaccharides, on the other hand, can be combined to form up to thirty-two distinct disaccharides when all possible combinations of linkage position (connection at C-2, 3, 4, 6 or 8) and anomericity (α or β configuration) are considered[1]. Common mammalian monosaccharides comprise glucose (Glc), mannose (Man), galactose (Gal), *N*-Acetylglucosamine (GlcNAc), fucose (Fuc) and sialic acid. Additionally, multiple residues can be attached to a single monosaccharide which can lead to complex branched structures. Therefore, oligosaccharides are amongst the most structurally diverse biopolymers in nature. The high structural complexity of oligosaccharides raises the need for introducing a symbolic representation and nomenclature for their convenient description. Figure 1 displays the drawing and abbreviation scheme based on Harvey *et al.*[2], which will be used throughout this work.

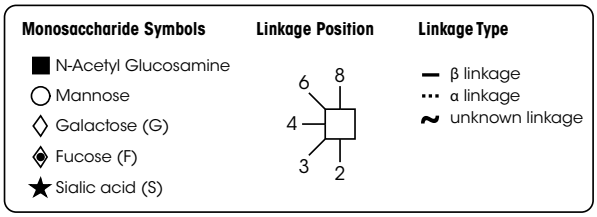



***Figure 1:*** *Symbolic representation and monosaccharide abbreviations. Based on*[2].

After or already during transcription, newly assembled poly-peptide chains are frequently structurally extended by co- and/or post-translational modifications. The transformation of proteins to glycoproteins by the attachment of sugars to the protein backbone is a highly common modification. It has been widely reported that almost all extracellular proteins are glycosylated to some extent[3] and more than 1% of the total human genome encodes for the approximately six hundred molecules involved in oligosaccharide biosynthesis[4]. Sugar moieties, also commonly referred to as glycans, can be attached to the nitrogen atom in the side-chain amino group of asparagine (*N*-linked type glycans) with a conserved consensus sequence of Asn-Xxx-Ser/Thr (where Xxx is any amino acid but proline) or to the oxygen atom in hydroxyl groups of threonine and serine (*O*-linked type glycans).

## *2.2 N*-Glycan biosynthesis is elaborate and built upon a multitude of enzymes

The biosynthesis of *N*-linked glycans in the endoplasmatic reticulum (ER) and Golgi apparatus (GA) of eukaryotic cells is an energetically elaborate process. A plethora of different highly specific enzymes promotes (i) the assembly of the lipid-linked precursor *N*-glycan, (ii) its subsequent trimming by certain monosaccharide residues in the ER and (iii) further attachment of new residues at non-reducing termini in the GA[5]. Figure 2 provides a schematic overview of the sequential enzymatic reactions involved from the attachment of the lipid-linked precursor to the poly-peptide backbone by the oligosaccharide transferase (OST) until the final maturation of the glycans in the *trans*-Golgi.

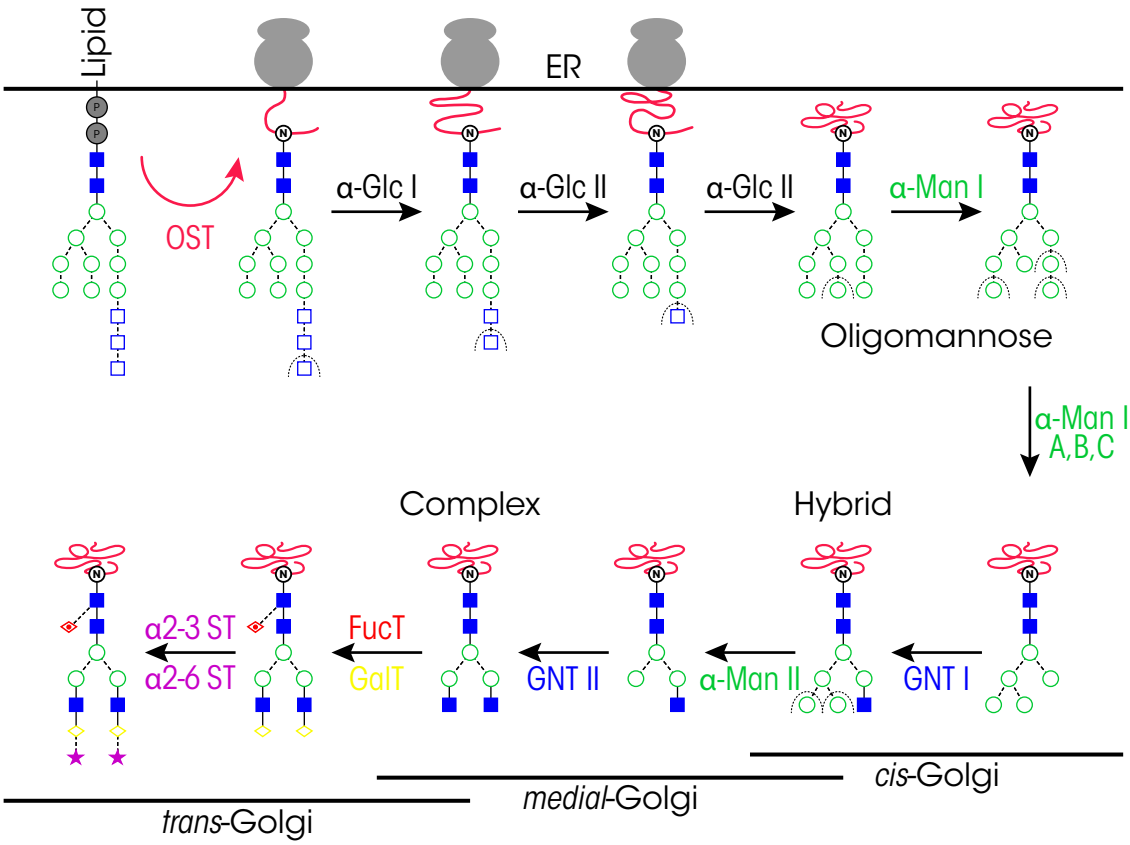



***Figure 2:*** *Schematic biosynthetic pathway of N-glycans in eukaryotic cells.*

Trimming of the 14-sugar precursor by exoglycosidase enzymes involves removal of the terminal glucose residues and some or all of the α1-2 linked mannose residues, *via* α-glucosidases (α-Glc I & II) and mannosidases (α-Man I, IA, IB, IC), respectively[6]. The resulting intermediate glycans comprise of five to nine mannoses attached to two GlcNAc residues of the chitobiose core structure ($Man_{5-9}GlcNAc_2$) and are commonly referred to as high-mannose type oligosaccharides. Some glycoproteins carrying high-mannose type glycans can escape further processing; however, more commonly $Man_5GlcNAc_2$ structures undergo further enzymatic processing. An additional GlcNAc residue is transferred to the α1-3 arm extending the chitobiose core *via N*-acetylglucosaminyl-transferase (GNT-I) activity and

terminal α1-3 and α1-6 linked mannoses are trimmed by the action of mannosidase II[5]. GNT-II catalyzes the addition of a GlcNAc to the previously pruned α1-6 arm, thus creates the basic structure for all complex type glycans[5]. Hybrid type glycans, comprising of a high-mannose and a complex type branch, result from mannosidase II inactivity. The basic complex type bi-antennary (A2) glycan can be further extended by the action of GNT-IV and V, adding a GlcNAc residue to respective α1-3 (C-4 position) and α1-6 (C-6 position) arms to form tri- (A3) and tetra-antennary (A4) glycans[6]. The enzyme GNT-III promotes the introduction of a β1-4 linked bisecting GlcNAc residue to the core mannose. Besides the bisecting residue, all other terminal GlcNAc residues of complex type glycans are commonly extended each by a galactose and a sialic acid residue, *via* respective activity of galactosyl- (GalT) and sialyltransferases (ST). Terminal sialic acids, most commonly *N*-acetylneuraminic acid (NeuAc) in mammals, can be either α2-3 or α2-6 linked to their adjacent galactose residues. Another common modification is the addition of a fucose residue by fucosyltransferases (FucT) to either the reducing terminal core or to one of the antennary GlcNAc residues.

## 2.3 Impact of sugar moieties on the parent protein

The attachment of sugar moieties can have pivotal influences on the parent proteins. Glycans not only alter physico-chemical properties such as molecular mass or charge state, but they can also impact folding[7-8], recognition factors with binding partners[9], immunogenicity modulation[10], regulation of bioactivity and final degradation[11]. The impact of glycosylation can be far-reaching, but particular effects are generally specific for the affected protein family. Therefore, immunoglobulin G (IgG), haptoglobin and lysosomal acid hydrolases have been chosen as highly interesting representative glycoproteins and the impact of their glycosylation will be described in further detail.

**Immunoglobulins** comprise of a highly abundant class of serum glycoproteins that provide a link between the selectivity of the adaptive immune system and the cellular responses of the innate immune system. IgG molecules are the most common of the serum antibodies, that consist of two light and two heavy chains. Each heavy chain possess a single covalently attached bi-antennary *N*-glycan at the highly conserved asparagine 297 residue in each of the $C_H2$ domains of the Fc portion of the molecule[12]. The attached oligosaccharides are structurally important for the stability of the antibody[13]. The *N*-glycans are also involved in the modulation of physiologically relevant effector functions[14]. Therefore, manipulation of the glycosylation of recombinant monoclonal antibodies to elicit optimized therapeutic effector functions represents an important research area[15-16].

Deletion of the oligosaccharides has been shown to cause a conformational change resulting in reduced or ablated binding to Fcγ receptors that bind to the Fc region close to the hinge[17]. The stabilizing effect of the *N*-glycans arises due to multiple non-covalent interactions between the oligosaccharides and the polypeptide chains of the $C_H2$ domain[18]. Investigations showed that on each chain these interactions occur between the elongated antennae that extend from the α1-6 linked mannose residues of the trimannosyl cores whilst the antennae on the α1-3 linked mannose residues extend into the interstitial space between the $C_H2$ domains[19]. Furthermore, the *N*-glycans between the $C_H2$ domains help to maintain an open horseshoe type conformation that facilitates favorable interactions with Fcγ receptors[19]. As each heavy chain bears a single *N*-glycan in the Fc region, symmetric or asymmetric pairing of the oligosaccharides confined within the interstitial space of the $C_H2$ domain is also important with regard to maintaining favorable stability and subsequent effector function[20].

The oligosaccharides at asparagine 297 also play an important role in the modulation of IgG activity. The absence of α1-6 linked fucose on the reducing terminal GlcNAc residue has been shown to dramatically increase the interaction with Fc receptor FcγRIIIA and the associated elicitation of antibody dependant cell cytotoxicity (ADCC)[21-23]. The binding is hindered when the IgG *N*-glycan is core fucosylated[21]. The affinity of the antibody receptor interaction is also mediated by the oligosaccharide at asparagine 162 on FcγRIIIA. Interactions with other activating receptors such as FcγRIIA are unaffected by the presence of the α1-6 linked fucose[24]. The presence or absence of antennary galactose residues has also been demonstrated to alter the activity of antibodies. Increased levels of agalactosyl IgG have been reported in autoimmune diseases suggesting that the absence of galactose may have a role in the pathogenicity of autoantibodies. *In vitro*, these agalactosyl antibodies were able to activate complement *via* the lectin pathway initiated following recognition by mannose binding lectin (MBL)[25-27]. Recently, it has been demonstrated *in vivo* that triggering of complement dependent cytotoxicity (CDC) by agalactosyl IgG occurs *via* interaction with activating Fc receptors[28]. Sialylation has also been implicated in regulating IgG activity. In contrast to other serum glycoproteins, IgG *N*-glycans displaying terminal sialic acid represent a minor component of the total *N*-glycan pool[26]. However, IgGs displaying terminal sialic acid have been shown to exhibit anti-inflammatory properties due to reduced antibody receptor interaction and up regulation of the regulatory FcγRIIB receptor on certain cell types[29-30].

**Haptoglobin**, is one of the important acute phase glycoproteins, and as most serum glycoproteins, is produced by hepatocytes. It is a positive acute phase protein (*i.e.*, its level is increased in inflammation) with the main function to capture and transport the intravascular free hemoglobin to cells[31]. It is composed of two heavy (β) and two light (α) chains or the oligomers of these, with four glycosylation

sites located on each beta-chain at Asn184, Asn207, Asn211 and Asn241[32]. Glycosylation, on average, represents approximately 19% of its total mass[33]. If not glycosylated, the ability of haptoglobin to bind hemoglobin is altered[34], therefore, glycosylation is critical for its function.

Lysosomes, present in the majority of eukaryotic cells, play an important role in the degradation and recycling of proteins, lipids, glycoconjugates and nucleic acids through the activity of the various **acid hydrolases** maintained within this acidic single-membrane bound cytoplasmic organelle[35]. These acid hydrolases enter the secretory pathway *via* the endoplasmic reticulum and the Golgi apparatus but are selectively shuttled to the lysosme rather than secreted *via* classical exocytosis[36]. This specialized lysosomal delivery is facilitated by the addition of a mannose-6-phosphate (M6P) residue to an *N*-linked oligosaccharide present on a newly formed hydrolase. This reaction takes place in the Golgi by the sequential action of UDP-*N*-acetyl glucosamine-lysosomal-enzyme *N*-acetylglucosamine-phosphotransferase (GlcNAc-P-T) and *N*-acetylglucosamine-1-phosphodiester *α*-*N*-acetylglucosaminidase[37-38]. Upon reaching the *trans*-Golgi, the M6P tagged hydrolase interacts with either a cation-dependent or -independent mannose-6-phosphate receptor (MPR), type I membrane glycoproteins that are members of the P-type lectin family. The MPR delivers the hydrolase to the lysosome where the M6P moiety is dephosphorylated *via* the action of acid phosphatase 5 (ACP5)[39-42].

## 2.4 Altered glycosylation can serve as disease specific marker

Despite the biosynthetic complexity, glycosylation under controlled physiological conditions or during bioprocessing is a highly reproducible and cell specific process. Therefore, changes observed in the glycosylation profile are indicative of aberrant changes in the physiological or pathological status of the cell[43-45]. Alterations in the glycome in the cancerous state have long since been reported[46]. The presence of cancer associated alterations in glycosylation offers the potential for the discovery of a new category of glycan based biomarkers.

Considerable research effort has been invested in the discovery of alterations in glycosylation offering potential diagnostic or prognostic monitoring capabilities. Biomarkers based upon increased sialylation and fucosylation of *N*-glycans present on serum glycoproteins have been reported in ovarian[47-49], breast[50-55], pancreatic[32, 56-58], prostate[59-61], liver[62-64], and lung[65] cancers. Although it appears that the reported markers are cancer associated rather than cancer specific, they nevertheless provide an insight into the molecular mechanisms of carcinogenesis and disease pathogenesis[66]. For example, Seales *et al.* have demonstrated the role of *α*2-6 linked terminal sialic acid, a commonly noted increase in the cancer serum glycome with increased metastatic potential of colorectal cancer cells[67]. In addition to

this, Abd Hamid *et al*. performed a longitudinal study to monitor breast cancer patients and reported that the proposed glycan biomarkers could detect a metastatic occurrence at a much earlier time point than the current protein marker, CA15-3[50].

Depletion of serum glycoproteins such as transferrin, haptoglobin, α1-acid glycoprotein or IgG and separate monitoring of their glycosylation changes in the course of disease progression can provide more detailed insights in the underlying molecular processes[56]. For example, understanding the role of altered IgG glycosylation in autoimmune disease and cancer is important to elucidate both the molecular pathology of such conditions and also the identification of possible targets for subsequent drug development[9, 68-71].

As another specific example, the addition of mannose-6-phosphate to acid hydrolases is a critical process in the targeting of enzymes to the lysosome. Malfunction of this system results in the accumulation within the cell of substrates of the acid hydrolases including glycogen (Pompe disease), mucopolysaccharides (Hurler and Hunter Syndromes) and glycosphingolipids (Gauche and Fabry's diseases). This results in numerous serious physiological conditions referred to as lysosomal storage diseases[72-73]. Enzyme replacement therapies (ERTs) have been successfully used for the treatment and management of lysosomal storage diseases[74-75]. A critical quality attribute of ERTs is the presence of M6P residues on the *N*-glycans of these recombinant therapeutic proteins. Increasing the levels of M6P on ERTs has been observed to promote uptake and subsequently the activity of treatments for lysosomal storage diseases both *in vitro* and *in vivo*[76-78].

## 2.5 Glycan structural elucidation poses a substantial analytical challenge

Structural elucidation of protein derived *N*-linked oligosaccharides represents a considerable analytical challenge due to the previously outlined inherent complexity and their branched nature. Both the absence of chromophores or fluorophores and the potential occurrence of very closely related structures (including positional and/or linkage isomers) greatly contribute to the challenges of oligosaccharide analysis and require high resolution techniques capable of comprehensively characterizing complex glycan mixtures. Furthermore, the same glycoproteins may be occupied by diversified carbohydrate moieties, even at the same glycosylation site (microheterogeneity), or exhibit different glycosylation site occupancy (macroheterogeneity).

Matrix assisted laser desorption ionization mass spectrometry (MALDI-MS) and high performance liquid chromatography of fluorescently labeled oligosaccharides (HPLC-fluorescence) have emerged as the dominant analytical techniques applied for the discovery of glycan based biomarkers in biological fluids or tissue samples[79-80]. To facilitate ionization and stabilize labile sialic acid residues, permethylation of

*N*-glycans is often a necessary sample preparation step prior to MALDI-MS. Novotny and co-workers have revolutionized this once time consuming and cumbersome protocol and successfully reported a high throughput permethylation procedure prior to mass spectrometric determination[81]. Despite these advances, mass spectrometric analysis of oligosaccharides is not without limitations, mainly due to an inability to differentiate between isobaric glycan structures without using tandem mass spectrometry and problems with quantitation arising from competitive ionization of different glycan structures within the ion source of the mass spectrometer[54]. Separation techniques such as HPLC and capillary electrophoresis hold an advantage over MS based analyses due to their ability to separate isomeric oligosaccharide structures along with more reliable quantitation[82-83].

## 2.6 Hydrophilic interaction liquid chromatography *N*-glycan separation

Hydrophilic interaction based stationary phases (HILIC) have found wide spread use for glycan separations since the pioneering work of Alpert *et al.*[84-85] and offer the advantage of predictive selectivity with retention increasing with increasing glycan size and hydrophilicity[47, 50-51, 55-56, 58]. In conjunction with high reproducibility, HILIC-based HPLC analysis of *N*-glycans and data interpretation became much more accessible but also reliable by the introduction of relational databases, associating elution positions with glycan structures[86]. Both the advent of such databases and high-throughput analytical methods[87] have consolidated the importance of hydrophilic interaction based *N*-glycan separation as a reliable and robust technique. However, a limitation of HILIC based analyses is the amount of time needed per chromatographic run, generally requiring 60 minutes or more.

Ultra performance liquid chromatography (UPLC) and the advent of sub-2 μm stationary phase technology represent a significant advance in the analytical sciences allowing for high efficiency separations and reduced analysis times[88-91]. UPLC has found widespread application particularly in the area of metabolomics wherein the high peak capacities of reversed-phase sub-2 μm stationary phases are often used in conjunction with the accurate mass capabilities of time of flight (ToF-MS) and Orbitrap mass spectrometry for the characterization of alterations in the metabolome to monitor the pharmacokinetic fate of molecules[92-94] and the metabonome in an attempt to further understand cellular alterations in disease[95-96]. Application to lipidomics has also been reported[97-98]. UPLC is also attracting attention as the focus shifts from discovery based disease relevant proteomics towards high throughput quantitation of the observed alterations in the proteome[99-101]. Until recently, UPLC has received limited interest in the field of glycomics due to the lack of appropriate sub-2 μm stationary phases. Of those studies reported, Inagaki and co-workers described the application of reversed-phase UPLC profiling of fluorescently derivatized asparagine linked *N*-glycans released from glycoproteins[102-103].

Jones *et al*., reported the use of ion pairing reversed-phase UPLC for the characterization of heparin and haparan sulphate disaccharides[104]. The recent introduction of tailored HILIC stationary phases for UPLC offers the prospect of fast, high efficiency separation of oligosaccharides[105].

## *2.7 N*-glycan analysis by capillary electrophoresis

Unlike liquid chromatographic approaches such as hydrophilic interaction, reverse-phase (RP) or porous graphitized carbon (PGC) chromatography for carbohydrate analysis wherein isocratic or gradient elution is required[106], capillary electrophoresis (CE) based separation is driven by electric field mediated migration thus offering different selectivity.

### 2.7.1 Electrophoretic separation principle

Zone electrophoresis is a differential migration based method enabling separation of species with different charge-to-mass ratios[107]. With the aid of a sieving matrix, the technique can also separate analytes with the same charge-to-mass ratio but different molecular masses[108]. Glycans in nature often form heterogeneous pools of both neutral and charged structures having diverse or, in the case of positional and/or linkage isomers of some oligosaccharides, identical molecular masses[109]. In particular, the separation of such carbohydrates, holding equal charge-to-mass ratios with the same molecular mass, represents a major analytical challenge for any kind of separation technique.
In zone electrophoresis the drag force ($F$) acts on a moving species against its electric field driven motion and it is proportional to its migration velocity ($v$) as well as to the Newtonian viscosity ($\eta$) of the background electrolyte[110]:

$$F = k \cdot \eta \cdot v \qquad \text{(eq.1)}$$

According to Stokes' law, for spherical particles the constant $k$ can be substituted with $6\pi r$, where $r$ is the hydrodynamic radius of the species. If the acceleration caused by the electric force is counterbalanced by the drag force, the species with an associated charge of $q$, moves with a constant migration velocity that is proportional to the electric field strength. The proportionality factor, referred to as electrophoretic mobility ($\mu_e$), is equal to the migration velocity divided by the electric field strength ($E$)[111]:

$$\mu_e = \frac{v}{E} = \frac{q}{6\pi\eta r} \qquad \text{(eq.2)}$$

Eq. 2 suggests that the separation principle in zone electrophoresis is rather based on differences in charge-to-hydrodynamic radius than in charge-to-mass ratios of the analytes[112]. Due to this separation

phenomenon, structural and linkage isomers of oligosaccharides carrying identical net charge-to-mass ratios can be resolved in capillary zone electrophoresis, based on mere differences in their hydrodynamic radii.

### 2.7.2 Technical overview

The technical simplicity of capillary electrophoresis can be deceptive; CE is a highly adaptive versatile technique, enabling several separation modes simply by altering separation conditions such as the background electrolyte used. Controlling the electroosmotic flow (EOF) phenomenon, that is the directional bulk flow generated under an electric field by ions attracted to capillary surface charges, and its direction (co- or counter-electrophoretic mobility, $\mu_e$), allows altering the migration time window, experimental run time and enables the motion of oppositely charged species towards the detector. EOF towards the detection site facilitates a stable liquid flow and CE current for electrospray ionization (ESI) with hyphenated mass spectrometric detection[113] or can generate a reverse migration order[114-115] when driving against analyte electrophoretic mobilities (with $\mu_{EOF} > \mu_e$). Adjustment or suppression of the EOF by covalent or dynamic coating of surface charges can remarkably increase experimental reproducibility. Under arheic (without flow) conditions, capillary zone electrophoresis (CZE) separates analytes based on their charge to hydrodynamic volume ratio differences thus rendering it a powerful tool for carbohydrate isomer identification. The selectivity of CE for the separation of similar oligosaccharides can be improved by introducing secondary equilibria, such as borate complexation[116], micellar surfactants[117-120], chromatographic (pseudo-) stationary phases[121] or by polymeric additives[122].

Protein derived carbohydrates from physiological samples are often only available in limited quantities, lack of chromophores or fluorophores and frequently contain differentially charged species. Therefore, derivatization with charged fluorogenic agents is a usual way to increase detector sensitivity and enable simultaneous separation of neutral and charged analytes using electric field mediated methods. While optical on-column detection techniques such as laser induced fluorescence (LIF) enable high sensitivity and selectivity they do not however, provide additional glycan structural information. The choice of labeling agent is strongly dependent on the application and needs to be carefully selected according to the background electrolyte (BGE), EOF, sample characteristics and also the detection method. Enzymatically released *N*-linked glycans by peptide-N4-(*N*-acetyl-β-glucosaminyl)asparagine amidase (PNGase F) maintain the free amino group from the side chain of the parent asparagine and the resulting glycosylamine can be derivatized with amine reactive dyes under basic conditions[123-124]. More commonly, the liberated glycosylamines are converted to reducing sugars (aldehyde form) at acidic pH

and reacted with primary amino group containing dyes *via* reductive amination[125-127]. However, low pH and high temperatures can promote analyte hydrolysis and the potential loss of labile glycan constituents as well as selective labeling need to be diligently precluded by optimizing the derivatization reaction[128-129]. A favorable derivatization agent not only needs to meet reaction chemistry, size and pH stable charge state requirements, but also has to be compatible with the available detection platforms (*e.g.* laser or LED excitation wavelength). On-column complexation[130] or charged residues (*e.g.* sialic acids) can also provide sufficient separation mobility. Underivatized glycans are unfavorable for sensitive optical detection but are commonly applied in mass spectrometric analysis[131]. A free reducing end enhances MS fragmentation options but may also increase analyte complexity due to anomericity[132]. Conjugated glycans potentially reach higher ionization yields[126, 133].

The transition to on-line mass detection techniques with electrospray ionization brings about the necessity to alter CE separation conditions, including replacement of outlet buffer reservoir and associated electrical circuit closing as well as introducing volatile background electrolytes, to achieve MS compatibility. Key for successful hyphenation of CE and MS are interfaces that produce a stable spray with low flow (preferably <20 nL/min) and allow nearly independent optimization of the separation and ionization sections. Although, a certain degree of compromise needs to be accepted: popular sheath flow interfaces use a sheath liquid at the capillary outlet, which can be optimized for stable ionization, but at the cost of sensitivity due to sample dilution and higher flow rates. Direct or sheathless connections generally incorporate low flow rates, high sensitivity and reduced ion suppression but a common BGE for reasonable separation performance and ionization yield needs to be found. Liquid junction interfaces introduce a liquid lined gap between the separation capillary and ionization source that allows decoupling optimization of CE separation and ESI parameters at low sample dilution effects. In-depth information about CE-MS technology, including glycan analysis using off-line interfacing and further ionization techniques such as matrix-assisted laser desorption/ionization (MALDI), can be obtained from eminent recent review articles[131, 134-136].

Miniaturization of CE into micro-channel devices offers another attractive ESI-MS front end due to high sensitivity analysis of nanoliter amounts of sample[137]. Irrespective of the chosen detection method, micro-channel glycan separations generally offer ultra-fast separation times on the seconds scale but usually at the cost of resolution; particularly of glycans carrying additional charged constituents, which are often removed prior to analysis[138-140]. For fast screening, *e.g.* of mAb glycans during clone selection or rapid profiling of clinical samples in disposable chips, where throughput is decisive, a lower separation efficiency can be acceptable. Moreover, recent introduction of longer separation channels has converged CE and micro-channel CE efficiencies but such designs yet need to be commercialized[141-143].

A dramatic gain in throughput is also possible by simply running multiple separation columns in parallel. Multiplexed systems are commercially available and incorporate from 4 up to 96 capillaries. High experimental precision is maintained by co-injection of internal standards for alignment. Multi-capillary systems are becoming increasingly established in clinical trials and biomarker discovery studies, where hundreds of samples are screened[144-145]. However, commonly built-in electrokinetic injection systems require careful sample handling, such as maintaining consistent salt or free dye content[146]. Avoiding selective analyte injection is of utmost importance when peak area based quantitation and associated statistical analyses are conducted.

### 2.7.3 Glycan structure elucidation

CE offers a variety of separation modes, which incorporate different selectivity and associated glycan structure identification potential. Counter-electroosmotic flow enables increasing charge based migration grouping with respective larger species migrating prior to smaller ones of the same net charge[147]. Selectivity enhancement of structurally related glycan pools can be achieved by online interaction with ionic buffer additives, such as borate[148], magnesium[149], or charged cyclodextrins[120], exhibiting different secondary equilibrium based upon monosaccharide composition. Zone electrophoresis in aqueous or low viscosity polymer matrix with suppressed EOF provides exceptionally high resolution of isomeric species due to the separation principle of charge-to-hydrodynamic volume differences[150].

**Oligosaccharide standards.** Co-injection of purified glycan standards with an unknown sample mixture poses a simple and effective way for structural assignment. Although, basic knowledge about the nature of the analyte pool present is useful, because glycans with different monosaccharide or linkage patterns can exhibit identical migration behavior. Limited commercial availability of structures, their production *via* purification with techniques offering restricted separation efficiency (*e.g.* isomers), and taxonomy mismatch with the sample further restrict practical implementation of standards.
Anticipation of molecular size from analyte migration time can be achieved by the ancillary separation of gradually sized oligomeric sugar standards. Optimal reference standards incorporate equally distributed oligosaccharide pools exhibiting a linear relationship between size and migration time. Thus, linear homooligosaccharide ladders with degrees of polymerization (DP) 1, 2, ..., n, of glucose(α1→4 or 6)n, glucose(β1→4)n and *N*-acetylglucosamine(β1→4)n, produced by hydrolyzing starch, cellulose and chitin-type polysaccharide chains, respectively, are commonly applied standards. Hydrolysates of branched structures, *e.g.* high-mannose type glycans[151], are generally not recommended due to the occurrence of positional isomers and the limited DP range. Molecular size

standards can either be co-injected with the sample, *e.g.* oligomer fragments of DNA base pairs in multi-capillary sequencer studies[152], or analyzed in a separate experimental run, where alignment standards are usually introduced in both sample and standard run[153-154]. Figure 3a and b show the separation of a glucose-oligomer ladder and a purified glycan standard, respectively.

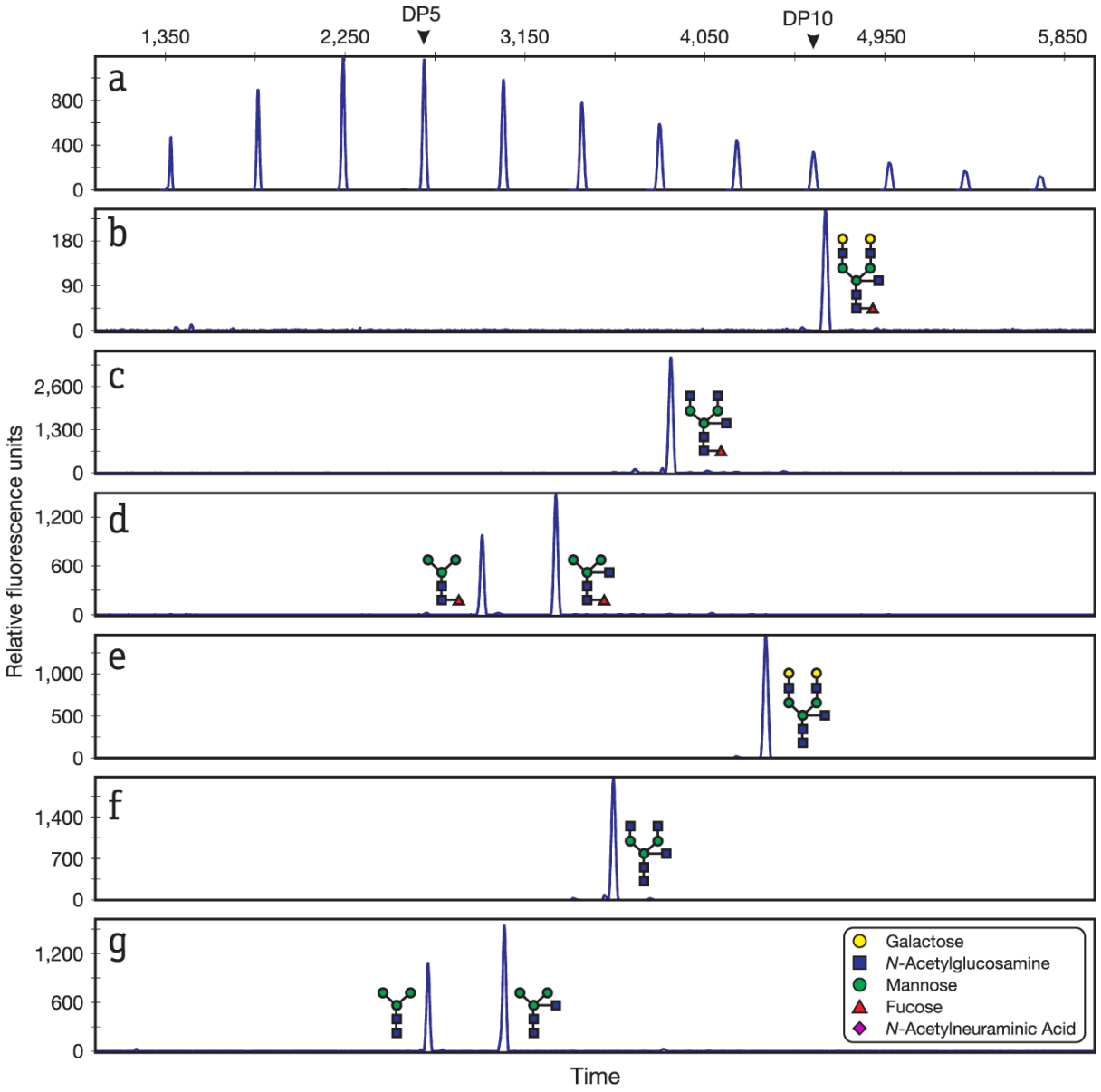



***Figure 3:*** *Exoglycosidase sequencing of a purified biantennary core fucosylated complex glycan standard with a bisecting N-acetylglucosamine. (a) Separation of an 8-aminopyrene-1,3,6-trisulphonic acid (APTS)-labeled oligomaltose hydrolysate reference standard. The number of glucose units (DP, degree of polymerization) in these structures is indicated. (b) Non-digested standard. Standard digested with (c) β-galactosidase, (d) β-N-acetylhexosaminidase, (e) α-fucosidase. Standard digested simultaneously with (f) β-galactosidase and α-fucosidase, (g) β-galactosidase, β-N-acetylhexosaminidase and α-fucosidase. Symbols used for glycans are those suggested by the Consortium for Functional Glycomics. Adapted from[155].*

The application of a lower (migrating faster than any sugar structure of interest), and an upper bracketing standard (migrating slower than any peak of interest) can elevate the experimental precision in CE[156]. Thus, the migration times of the peaks within the boundaries can be converted to relative migration times (RMT) using the bracketing standards as follows:

$$RMT = \frac{t - t_{lower}}{t_{upper} - t_{lower}} \quad \text{(eq.3)}$$

Eq. 3 normalizes migration times into the interval between 0 and 1, where $t$ is the migration time of the analyte, $t_{lower}$ and $t_{upper}$ are the respective migration times of the lower and the upper bracketing standards. Co-injection of the bracketing standards with the sample allows for such horizontal alignment, thus significantly increases analytical precision, not only between consecutive repetitive runs but also between oligosaccharide ladder standard and sample runs.

The respective DP or sugar units (SU) of a sample peak can be obtained by interpolation between adjacent oligosaccharide ladder standard peaks, based on the well-established Kovats retention index in gas chromatography[157]:

$$SU_x = S_n + \frac{t_x - t_n}{t_{n+1} - t_n} \quad \text{(eq.4)}$$

where $SU_x$ is the sugar unit value (corresponding to DP) of the unknown sample peak, $S_n$ is the DP of the preceding oligosaccharide ladder peak, $t_x$ is the migration time of the unknown sample peak, while $t_n$ and $t_{n+1}$ are the migration times of the homooligomers immediately preceding and following the peak of interest, respectively.

The conversion from migration time to a size-based scale of SU also promotes inter-experiment, -instrument and -institution precision and comparability, by compensating for potential buffer composition, separation temperature or column history induced experimental deviations. The accumulation of SU values of glycans with known (*e.g.* purified standards) or decoded structural identity can evolve into the establishment of a database, similar to the existing one for HILIC[86], capable of decisively supporting structural elucidation or *de-novo* sequencing of unknown glycan pools by CE.

**Enzymatic digests.** The specific cleavage of monosaccharide constituents from non-reducing termini *via* exoglycosidase enzymatic digestion represents another powerful means for glycan structure characterization. Depending on enzyme specificity, monosaccharide type, sequence, or even linkage and anomericity can be obtained when digest induced structural/shape changes are accompanied by CE monitoring. Sophisticated digest cascades can also lead to the identification of positional isomers[158-159].

Exoglycosidase digest arrays can either be conducted in parallel, where the sample is evenly split into one aliquot per enzyme mixture, or in sequential fashion using analysis-digest iterations of the same sample. Sequential processing requires lower amounts of sample and enzymes and takes advantage of the minimal injection volumes needed for CE analysis. The use of non-volatile buffer components or addition of alignment standards to the sample prior to injection, as for example used in multi-capillary sequencer experiments, can impede further digestion steps. Moreover, exoglycosidase digests can cause changes in sample ionic strength or pH, introduced by salts of the digestion buffer, and promote diffusion induced peak distortion as well as potentially altered electrokinetic injection behavior[160]. Volatile digestion buffer systems are therefore generally recommended to alleviate this issue. Most commercial enzymes are, however, supplied in non-volatile media and consequently will cause the described problems in CE, if concentrated sample volumes or multiple enzymes are used.

Exoglycosidase digestion of glycans with known structure (*e.g.* purified standards) can cause a predictable loss of constituents corresponding to enzyme specificity. After reanalyzing the digested product by CE, glycan structural differences can be recorded by the differences in migration time and relative peak area compared to the substrate. Such shifts in migration time, or respective sugar units, can be used to build a knowledge base of monosaccharide residues and associated 'contributions'[161]. Traces c-d and e-g in Figure 3 exemplify two different digest cascade possibilities of the purified glycan standard analyzed in trace b. Structural identity of hydrolyzed products, respective SU, as well as digest induced shifts can readily be deduced. Additionally, positions of unknowns, *e.g.* a monogalactosylated biantennary species, can be anticipated half way between the bi- and the agalactosylated peaks as respectively depicted in Figure 3b and c.

In case of dealing with unknown glycan mixtures, increasing structural identity confidence can be gained by tracing respective peaks through several digestion steps, or by virtual 'reattachment' of constituents in a bottom-up manner, especially when combined with sugar unit shift analysis and CE-based glycan structure reference databases. Relative peak areas before and after digestion nevertheless need to be rationally compared when analyzing complex oligosaccharide pools. Glycans composed of different monosaccharide units can exhibit identical migration, due to *e.g.* similar charge to hydrodynamic volume properties, but differentially or even unaffected enzymatic digest reactivity. This can be either due to lacking the epitope that matches enzyme substrate specificity, or inaccessibility potentially induced by steric hindrance. For example the removal of bisecting *N*-acetylglucosamine (GlcNAc) residues upon hexosaminidase treatment can be hampered as depicted in Figure 3d and g.

The hydrolysis of sialic acids, which introduce additional molecular charges, by sialidase treatment prior to analysis, is a rather regularly applied practice in CE based glycan investigations[140, 144-145, 162-165].

However, associating sialic acid removal with technical limitations of electromigration based methods is a misapprehension. For example, CZE with suppressed EOF offers increased separation efficiency of additionally charged thus faster migrating glycans, due to decreased analyte diffusion. The removal of sialic acids generally results in decisive reduction of CE profile complexity, by merging previously distributed corresponding species with differential degree of sialylation, as well as sialic acid linkage and positional isomers. This reduction of complexity coincides with the loss of information and potentially correlated physiological features, when investigating glycans of biomedical interest[166-167]. Although, the combination of previously distributed low abundant species upon sialidase digestion might also allow for the investigation of other biologically important glycan features, formerly undetected due to dynamic range issues. When electrokinetic injection is used, equalizing analyte charge states will also support diminishing potentially biased injection.

Different endoglycosidases, used for the release of asparagine-linked sugar moieties from glycoproteins, can be applied for specific glycan type-based partitioning. While PNGase F liberates all classes of *N*-linked glycans from mammalian glycoproteins, Endo-β-*N*-acetylglucosaminidase H (Endo H) specifically releases high-mannose and hybrid type oligosaccharides[168]. Endo H cleaves substrate glycans between the two GlcNAc residues in the core region and readily enables the identification of the specific glycan types by CE[169]. Determination of the intersection and difference sets with the respective PNGase F released pool enables type-based classification[170-171], when the alterations by the missing GlcNAc residue and associated SU shifts are accounted for.

Exoglycosidase digests are usually carried out in low reaction volumes (~10 µL) with overnight incubation to achieve exhaustive enzymatic processing. On-column enzymatic digestions represent a practical rapid, low sample and enzyme consuming alternative with incubation times down to only a few minutes, or even during the separation process itself by passing the injected sample through an enzyme plug[172]. Direct introduction of exoglycosidases into narrow bore separation channels can increase enzymatic reaction speed due to decreased diffusion limitations but such endeavors are sometimes incompatible with certain enzymes[172] and often accompanied by some loss in separation efficiency[164, 173-174].

**Lectin affinity.** As opposed to consecutively cutting glycan constituents, carbohydrate binding proteins, incorporated into the separation column, can result in specific peak retardation or disappearance, based upon their respective structural affinity. Lectin affinity CE enables the structural classification of glycans by type (*e.g.* high-mannose[164, 174-175]), antennary branching (*e.g.* tri-antennary[176-177]), monosaccharide features (*e.g.* fucose[172, 177-178], bisecting GlcNAc[176], galactose[172], sialic acid[177]), or even by their glycosidic linkage type (*e.g.* α2-3 or 2-6 linked sialic acids[172, 179]), *via* specific binding reactions.

Similar to on-column enzymatic digestions, lectins can either be added to the BGE[175-179] or introduced as a distinct zone[164, 172, 174], but generally no incubation time is needed. Although, not all lectin-carbohydrate interactions are well understood, binding associations are generally low and peak distortion can limit area based comparison of experiments with/without lectins, affinity CE represents another fast and valuable structure identification tool.

**Additional separation dimensions.** Additional orthogonal separation dimensions offering different selectivity can drastically increase the resolution. Especially when investigating the identity of heterogeneous carbohydrate pools, the reduction of sample complexity by preparative separation and subsequent analysis of the collected fractions by CE is often indispensable. Analytes can either be partitioned into groups of multiple species according to sample properties such as charge state, or, often by using several preparative techniques, into distinct glycan species. The preparation of such purified carbohydrate standards enables the identification strategies outlined under **Oligosaccharide standards** and **Enzymatic digests**, when combined with exoglycosidase digestion. For highly complex samples consisting of a multitude of species, this approach can be extremely labor-intensive or even impossible due to restricted sample availability and resolution limitations of preparative techniques in separating closely related structures. Difficulties in structural identification arising from co-migration of species with similar charge to mass (or equivalent hydrodynamic volume) ratios in CE can be minimized by molecular charge or size based fractionation. Orthogonal separation mechanisms, such as analyte polarity based surface interaction, can also distinguish previously co-migrating species, due to differential separation selectivity.
Chromatographic techniques with various stationary phases and corresponding distinct selectivities accompanied by higher associated injection volume requirements are ideally suited for preparative fractionation prior to CE analysis. Compatibility with subsequent CE analysis is reached by desalting or the application of volatile liquid phases. Similar to glycan analysis by CE, sensitive detection in liquid chromatographic (LC) techniques often requires the introduction of fluorescent dyes. However, based upon the different selectivity of chromatographic phases, difficulties can arise from certain properties of labeling agents (*e.g.* charge, polarity) which were optimized for glycan separation by CE. The combined application of preparative LC and analytical CE techniques with sensitive detection thus generally requires a compromise.

*Anion-exchange chromatography* (AEC) fractionation is one of the popular means to separate glycan pools by their degree of sialylation or other charge inducing modification (*e.g.* phosphorylation). Each charged fraction may contain several glycan structures, differing *e.g.* in branching

degree, composition of monosaccharides and their position or linkage, but all exhibiting the respective equal net negative charge. In CZE under arheic conditions for example, co-migration of larger structures holding higher charge states with smaller analytes, can aggravate structural identification. Additional charges introduced by sialic acids can be removed *via* sialidase digestion and the resulting changes in relative peak areas monitored, as outlined under **Enzymatic digests**. In cases where the ratio of charged *vs.* neutral species is particularly unbalanced in favor of the latter, *e.g.* sialo- *vs.* asialo-glycans on human IgG, and only trace amounts of certain charged species are observed, exoglycosidase digestion induced increase of the corresponding neutral species can be marginal. Partitioning into charged fractions resolves overlaps between differentially charged species and enables the monitoring of peak shifts to unoccupied regions upon charge removal and direct comparison to the associated neutral fraction.

Figure 4 shows the high pH high performance AEC (HPAEC) fractionation and consecutive CE analyses of human serum *N*-glycans released *via* endo H. While sialidase digest of the total pool in trace BS could result in inconclusive changes of the corresponding neutral species, individual analysis of preparatively captured fractions and their sialidase digestions, revealed the identity of the underlying monosialylated hybrid type species, as presented in Figure 4B[169].

Highly charged labeling agents, such as the commonly used 8-aminopyrene-1,3,6-trisulphonic acid (APTS), can critically increase the total negative charge of analytes, thus potentially causing excessive retention on AEC phases. Moreover, the fundamental separation principle in AEC is based upon differences in charge distribution across the hydrated molecule. The subdivision into distant non- and reducing terminal charges further complicates the application of highly charged dyes for AEC-based fractionation. Using APTS as an ionic dye for CE separation and sensitive detection is still possible if fractionation is carried out using underivatized glycans and only subsequent dye conjugation. This was demonstrated by pulsed amperometric detection (PAD) following HPAEC separation[169]. Based upon its separation principle, AEC allowed for additional size-based separation within each charged fraction, when operated with fluorescently labeled glycans and sensitive optical detection[180]. Supplemental size-based fraction collection of underivatized glycans was also achieved by UV detection[181-182], or by HPAEC-PAD, which enabled sensitive elution monitoring at maintaining adequate resolution, as exemplary shown in Figure 4A. HPAEC-based separation generally requires a subsequent buffer exchange step due to non-volatile strongly alkaline mobile phases, the described associated ionic strength mismatch issues in CE separation and high pH-induced sample epimerization[183].

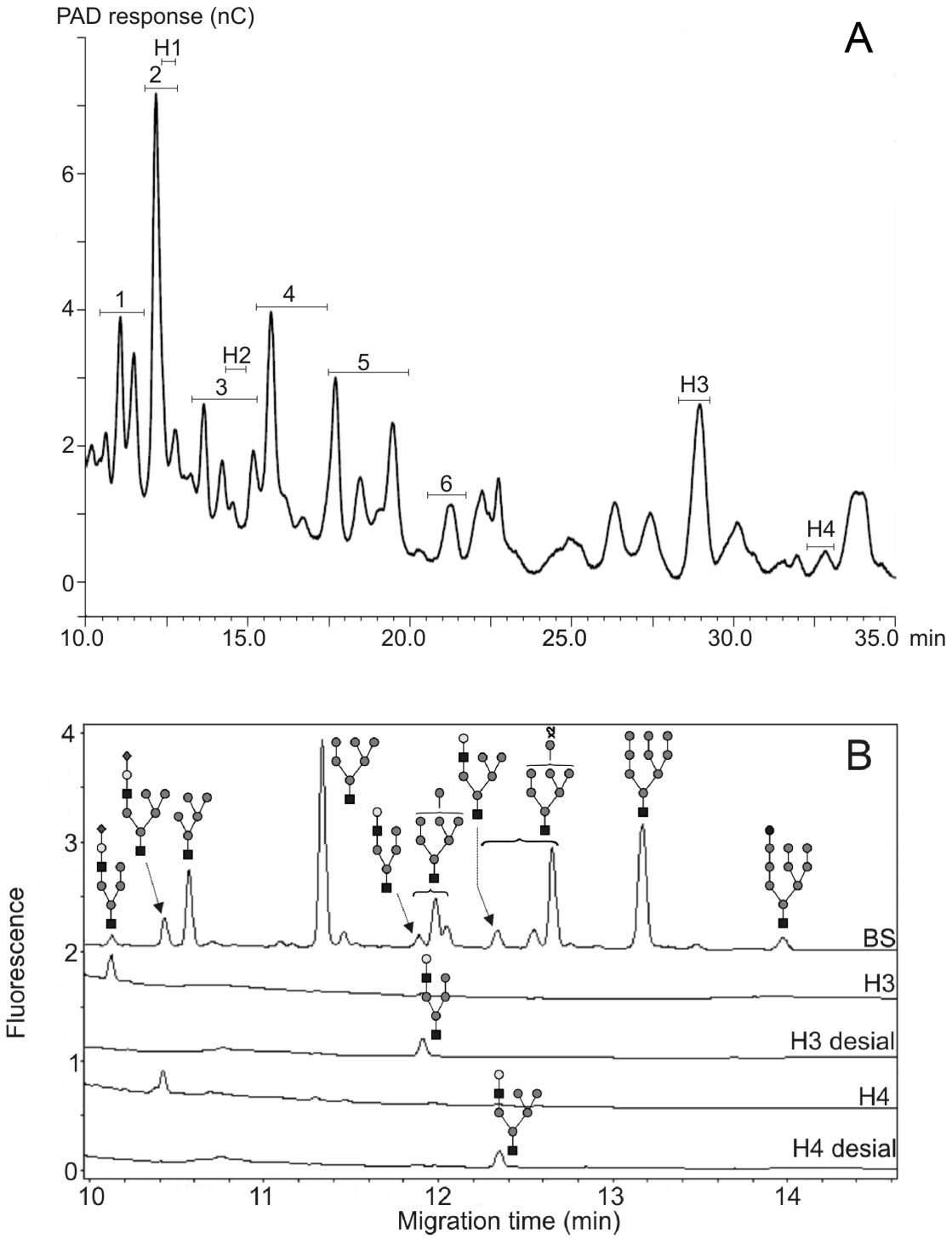



***Figure 4:*** *(A) HPAEC-PAD profile of human serum N-glycans released via endo H. Areas subjected to fraction collection are indicated. (B) CE-laser induced fluorescence detection (LIF) trace of APTS labeled total glycan pool from blood serum (BS), collected fractions H3, H4 and their respective sialidase digests (H3 desial, H4 desial). Glycan symbols as in Figure 3. Adapted from*[169].

*Hydrophilic interaction liquid chromatography* is a high resolution fractionation alternative to AEC, also offering orthogonal selectivity. Depending on the chromatographic column, retention is a function of hydrophilicity and associated glycan size, or a combination with ionic interactions[184], when using amide- or amine-based stationary phases, respectively. HILIC-based separation enables the collection of fractions with high individual structure purity as opposed to common analyte property based grouping in AEC, however, the higher resolution of the HILIC fractionation results in an increased number of fractions and associated larger scale downstream analysis. Combination of HILIC fractionation and exoglycosidase digestion of purified analytes is also commonly applied for accurate glycan structural elucidation[185-187].
Similar to AEC, the application of polar and ionic labeling agents such as APTS (LogP -1.21) for CE analysis, was preceded by either intact or removed charged glycan constituents, using underivatized[146] or APTS labeled oligosaccharides[155, 186], respectively. The motivations for these distinct strategies, however, seem to originate from mass spectrometric detection and CE instrument compatibility considerations. Other fluorescent labeling agents, such as 2-aminobenzoic acid, (LogP 0.78) featuring lower charge and polarity, were successfully applied for both HILIC-based fractionation and CE analysis using high sensitivity fluorescence detection[185, 188].

Applications of *reverse phase* (RP) and *size-exclusion chromatography* (SEC) glycan fractionation are limited due to lower associated applicability and selectivity, respectively, when compared to HILIC and generally low efficiency for relatively small sugars (0.1-5 kDa range), respectively. Only few high performance columns serving ranges in the low molecular weight region exist[189] and exclusive size-based preparative partitioning on ion-exchange columns suffers from low resolution[190]. As opposed to HILIC, retention on RP stationary phases is based on hydrophobicity thus providing only weak interaction for polar glycans. Although, type based separation of high mannose, complex and fucosylated complex glycans can be achieved[191] and RP-based fractionation could resolve co-migration of species from the distinct groups[192]. Also the derivatization with more hydrophobic dyes, such as 2-aminoacridone (LogP 2.95), can facilitate RP fractionation potential[193].

Capillary electrophoresis can also be operated in preparative mode using larger capillary diameters and several interchangeable outlet buffer reservoirs for fraction collection[194]. Low injection plug volumes often require several iterations of the fractionation process to achieve sufficient amount of analyte per collected fraction[195]. Micropreparative CE has nevertheless been successfully applied for off-line MALDI ionization and hyphenated mass detection, where higher concentration is reached by liquid phase evaporation and concentration within the MALDI matrix[196-197].

One often perceived limitation of CE is the low sample injection volume. However, this facilitates experiment repetition and also the consecutive execution of different CE separation modes. Hence, the simplification of complex samples *via* elaborate preparative fractionation can be avoided by the introduction of additional analytical separation modes with different selectivity. Glycans co-migrating in one separation dimension can be separated in another with appropriate selectivity, and *vice versa*. Such multidimensional mapping of oligosaccharides originates from the application of different stationary phases for liquid chromatography[198]. Nanoliter injection volumes and the versatility of CE allow the facile operation of several different separation modes, including EOF and on-column complexation phenomena, on the same instrument or even capillary, by simply changing the electrolyte used. Normalized migration times of identified carbohydrate species from each separation dimension span a multidimensional database thus facilitating structural identification of unknowns at increasing confidence[199-200].

Combination of analytical CE and LC glycan separation methods comes at costs of additional instrumentation and sample preparation steps, *e.g.* introduced by different fluorescent labels optimized for each technique, but enables orthogonal glycan separation[187]. Although, fluorescent dyes suitable for both platforms have been presented[83, 185, 188, 201], instrumental limitations such as available detection method, can restrict their application. Disregarding the increased necessary sample handling, the application of two distinct analysis routes allows for monitoring potential experimental discrepancies originated from *e.g.* loss of labile glycan constituents or salt content. Also one might not want to change a certain analysis route due to already existing databases that can be used for peak identification in the other dimension[181].

**CE-MS: Molecular mass and fragmentation.** Mass spectrometry offers an information-rich complementary detection method for glycan analysis, facilitating the measurement of distinct analyte mass and charge properties when used online, independent from standard co-injection or exoglycosidase digest patterns. Glycan composition can be estimated based upon the addition of monosaccharide constituent masses. However, due to the high structural diversity, numerous combinations can result in equal mass values. Moreover, on the monosaccharide level stereosiomers (*e.g.* glucose and mannose) exhibit the same masses, thus MS can only detect different classes of sugars such as hexose (glucose, galactose, mannose), *N*-acetylhexosamine (GlcNAc or GalNAc), deoxyhexose (fucose) or sialic acids. Information about the identity of monosaccharides may be deduced from additional sample information such as glycan class (*e.g. N*-, *O*-glycan), taxonomy and associated biosynthetic pathways. Such information, if available, provides very limited potential of elucidating linkage and positional isobaric

isomers. Therefore to increase the overall level of information in each associated experiment, MS detection can be combined with monosaccharide sequencing by exoglycosidase digest induced mass shift monitoring, or more commonly the application of tandem mass spectrometry (MS/MS) following collision induced dissociation (CID)[202]. In tandem MS or higher order $MS^n$ approaches of pseudomolecular oligosaccharide cations formed during positive ionization, glycans first break at their most labile bonds resulting in series of glycosidic cleavages. Glycosidic linkage cleavage allows determination of monosaccharide sequence and branching degree, based on B/Y or C/Z-ion series if negative ionization is used[203], whereas cross-ring cleavages and associated A/X-ions allow deduction of linkage positions[202]. Due to complexity and potential interpretation ambiguity, it is often reasonable to investigate fragmentation patterns with purified glycan standards similar to those under study. However, informatics platforms such as GlycoWorkBench[204] and GlycoPeakFinder[205] have greatly helped in simplifying the interpretation of oligosaccharide MS/MS spectra.

Problems arise from direct infusion of complex sample mixtures, where selected precursor ion masses potentially contain structural isomers. Fragmentation of multiple species can lead to incorrect spectral interpretation and glycan mass based structural conclusions. Hyphenation of MS with separation based techniques can provide additional structural information or even resolve isomeric species thus prevent erroneous structural interpretation from the resulting spectral data. Concentration of complex analyte pools into distinct bands using separation techniques results in increased dynamic range and associated sensitivity for MS detection. Furthermore, less analyte will enter the MS together and thereby reduce the suppression of less abundant ions by predominant ones. Ion suppression effects, as well as biased quantitation of differentially charged analytes and associated ionization yield can be resolved by upstream CE separation due to the inherently high efficiency of CE based separations.

MS interfacing generally comes at the compromise in CE resolution, introduced by a MS friendly BGE, the missing outlet buffer reservoir, differential ionization interface gas pressure, siphoning effects and band broadening in longer capillaries. On the other hand, the more informative MS detection facilitates the detection of co-migrating non-isobaric analytes. High resolution off-line CE glycan separation with photometric detection can be altered stepwise towards MS compatible conditions. Based upon relative abundances, respective peaks can be assigned in the CE-MS base peak electropherogram (BPE) and previously gathered structural information matched or even amended[183, 206-207]. The more accurate quantitation achieved by optical detection of heterogeneous analyte pools can also be incorporated online proximal to the mass spectrometer[208-210].

The separation of uncharged glycans in CE is problematic because on-column complexation with borate or metal ions is generally not MS compatible and sole EOF based separation lacks selectivity. Isocratic elution in CEC[132, 211] or the introduction of ionic labeling agents, enable the parallel separation of both neutral and charged species. Labels are chosen not only based on CE separation optimization, but also MS ionization considerations. Neutral or positively charged labeling agents, such as 9-fluorenylmethyl and 3-aminopyrazole, are used for positive MS ionization polarity modes[123, 212] and *vice versa* for the more commonly used APTS, 8-aminonaphtalene-1,3,6-trisulfonic acid and 2-aminobenzoic acid in the negative mode[83, 213-214].

Figure 5A shows the CE-MS BPE of APTS labeled *N*-glycans from the structural subunit 1 of Rapana venosa hemocyanin at m/z 555.7 isolated in Q1 of a quadrupole ion trap MS[212]. Counter-EOF conditions with the MS at the cathodic side allowed separation based on increasing negative charge and larger species migrating prior to smaller ones within each charged group. The peak at 9.5 min corresponded to the labeled $Man_5GlcNAc_2$ structure, whereas peaks at 10.5 and 10.9 min represent isomers of a different glycan species with an additional negative charge. Similar tandem MS spectra of the peaks at 10.5 and 10.9, depicted in respective Figure 5B and C, endorsed a matching monosaccharide sequence and suggested the presence of positional isomers that were separated due to the high resolving power of CE.

Labeling at the reducing end can provide valuable information on the site-attachment of specific glycan constituents due to fragment ions either including or lacking the mass and charge introduced by the labeling agent. A- and X-type ions and associated linkage position elucidation, however, are generally low when the reducing end of glycans is occupied, because cross-ring fragmentation predominantly occurs at the proximal end[215].

Unlabeled glycans that carry charged constituents, such as sialic acids or phosphate residues, were well separated in counter-EOF CZE based upon their charge degree and molecular size[216-219]. Such acidic glycans are most sensitively detected by negative ion MS, but positive mode allows quantitation of differentially charged species, *i.e.* glycan pools containing both positive and negative oligosaccharides, using *e.g.* acidic mobile phases in CEC that can protonate sialic acids[211]. Glycans with an unoccupied reducing end produce more informative MS/MS spectra due to higher abundance of A/X-ions, but reducing terminal mutarotation can increase sample complexity[132]. MS/MS fragmentation analysis of glycans holding terminal sialic acids often leads to predominant B1 ions originating from the loss of such charged glycan constituents and lower amounts of more informative C-type fragments[220].

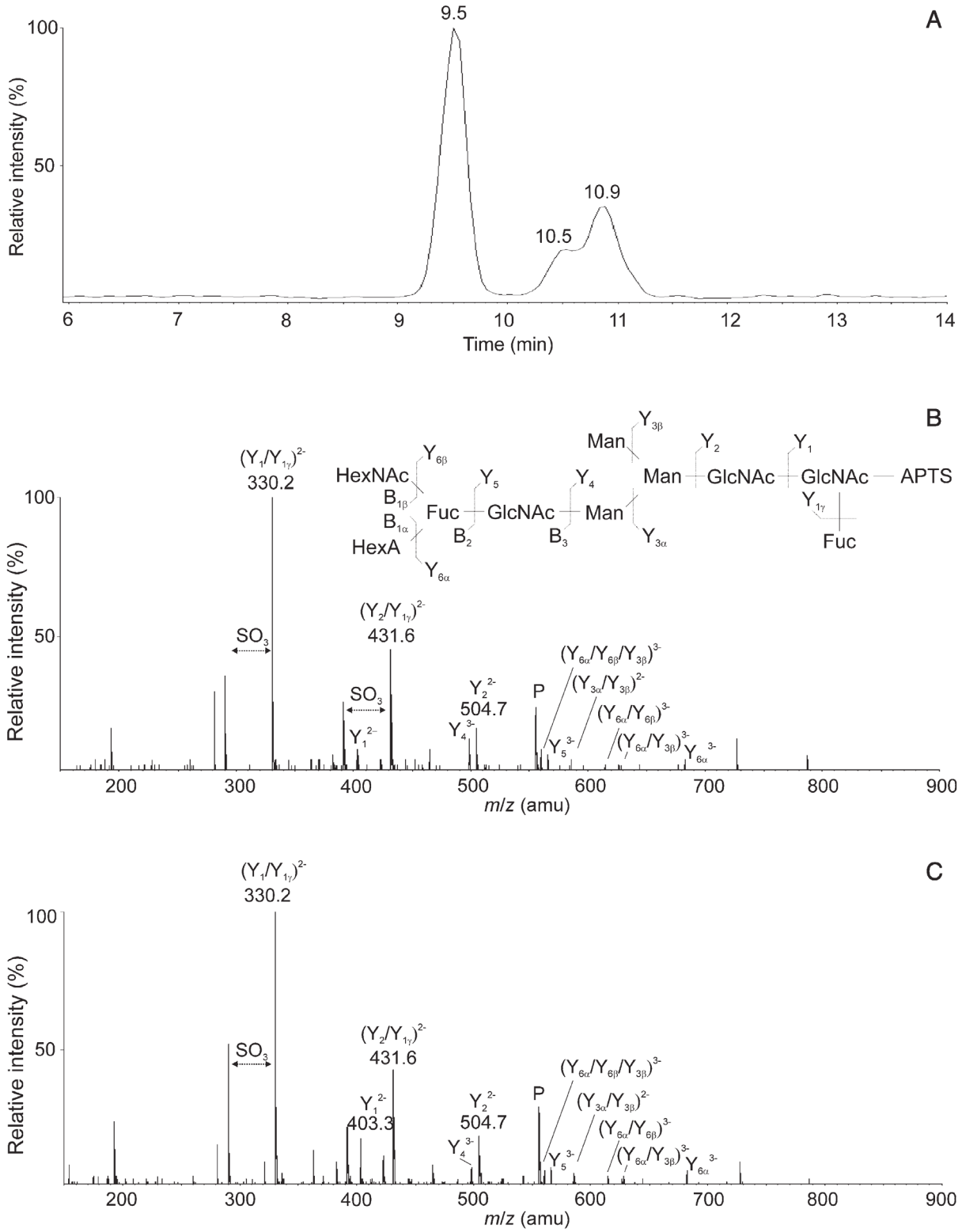



***Figure 5:*** *CE-MS/MS BPE of the fixed m/z 555.7 in Q1 of a quadrupole ion trap MS (A) and MS/MS fragmentation spectra of the respective peaks at 10.5 (B) and 10.9 min (C). The peak at 9.5 min corresponds to* $Man_5GlcNAc_2$ *at m/z 557.8. P, precursor ion. Fragmentation nomenclature as in*[203]*. Adapted from*[212]*.*

Chemical derivatization *via* amidation, methylation or permethylation can stabilize sialic acids[221] thus provide greater structural elucidation potential in MS/MS fragmentation[222], but coinciding charge neutralization and increased hydrophobicity limit their online CE-MS application. Offline CE fraction collection and spotting on MALDI plates combines CE efficiency with less complex MS spectra originated from mostly singly charged ions[196-197, 223]. Low mass loading in CE limits chemical derivatization options of collected fractions. These derivatizations are generally recommended to prevent glycan degradation during the higher energy laser desorption/ionization[224].

## 2.8 Aims of the study

Alterations in glycosylation of serum proteins have been noted in a number of diseases including cancer. HILIC proved valuable for structural identification and peak area based quantitation of serum glycans due to robustness, reproducibility and the availability of databases that aid elution time based structural annotation. A limitation of HILIC-HPLC based analyses is the generally long experimental run time. Therefore, the aims were to:

- Transfer the established HILIC methods for *N*-glycan analysis to sub-2 μm HILIC phases for UPLC to increase efficiency and selectivity at decreased analysis time.
- Demonstrate the importance of separating glycan isomers of potentially associated biological significance at higher efficiency and highlight the method applicability in identifying disease associated glycan changes.

CE represents a rapid alternative analytical technique, offering electric field mediated separation selectivity with the ability of separating isomeric species based on molecular shape differences. However, robust and reproducible endeavors with associated extensive mapping databases that enable confident structural inference from normalized detection times of fluorescently derivatized glycans have not been introduced. Thus, the goals were to:

- Develop a practical analytical method including eligible normalization approaches to ensure experimental precision as well as structural elucidation accuracy, matching those of the well established HILIC and the introduced improved UPLC approaches.
- Highlight the applicability of the method and the exceptional resolution of CE by the investigation of disease related *N*-glycan alterations.

Both UPLC and CE methods offer exceptional resolution of closely related isomeric glycan species and short analysis times. Although, the two techniques exhibited different selectivities based upon their underlying separation principles. Consequently, the objectives were to:

- Compare the two analytical methods and highlight their inherent merits and shortcomings.
- Investigate the potential of a combined parallel application to further decrease experimental expenditure but in the same time increase glycan structural elucidation accuracy and confidence.

# CHAPTER 3

# Materials and Methods

## 3.1 Chemicals and reagents

Reagent water used throughout this study was obtained from a Milli-Q Gradient A10 Elix system (Millipore, Bedford, MA, USA) and was 18.2 MΩ or greater with a total organic carbon (TOC) content less than 5 parts per billion (ppb). Solvents, LC-MS Optima grade or Chromasolv HPLC grade, were obtained from Thermo Fisher Scientific (Dublin, Ireland) and Sigma Aldrich (St. Louis, MO, USA), respectively. Purified recombinant β-glucuronidase, expressed in CHO cells, was provided by Prof. Jeffry H. Grubb. The protein, present in a buffer containing 25 mM Tris-HCl pH 7.5, 1mM β-glycerophosphate, 0.15 M sodium chloride and 0.025% w/v sodium azide, was aliquoted upon receipt and stored at -80°C until further use. Protein purity was assessed *via* reducing 10% SDS-PAGE and subsequent LC-MS/MS analysis following tryptic digestion. Purified *N*-linked carbohydrate standards and exoglycosidase enzymes were obtained from Prozyme (Hayward, CA, USA). Maltooligosaccharides were obtained from Grain Processing Corporation (Muscatine, IA) and chitooligosaccharides were from Seikagaku Biobusiness (Tokyo, Japan). All other chemicals used were purchased from Sigma Aldrich and were of the highest available quality.

## 3.2 Serum samples and small scale 2-DE for stomach cancer study

Blood samples were collected from 80 patients with stomach cancer, (18 Level I, 26 Level II, 23 Level III and 13 Level IV classified according to American Joint Committee on Cancer staging system, 6$^{th}$ edition) and 10 patients with benign stomach disease along with 20 matched controls at the Clinical Institute Fundeni (Bucharest, Romania) with informed consent by RNTech as part of the EU FP6 GLYFDIS research consortium. Venous blood was collected into Vacuette™ serum tubes and allowed to clot for approximately 30 minutes. Tubes were then centrifuged at 3,000 rpm for 5 minutes at room temperature. The separated serum was aliquoted, transported and stored at -80°C until further use. For the current study 5 μL of serum from each patient was pooled according to their pathological classification using the supplied clinical information to yield disease specific patient pools. Pooled samples were used in order to minimize biological variance. Detailed

information regarding the age, sex and pathological classification of each patient is provided in Supplementary Information Table S1 of[69].

A volume of pooled serum corresponding to 150 μg of total protein was diluted in rehydration buffer (8 M urea, 0.5% w/v CHAPS, 0.2% w/v DTT, 0.2% v/v Pharmalyte) and loaded onto 7 cm Immobiline DryStrips pH 4-7 *via* overnight passive rehydration (GE Healthcare, Upsala, Sweden). Isoelectric focusing was performed using a GE Ettan IPGphor 3 IEF system. Once focused, the strips were equilibrated in reducing buffer (6 M urea, 50 mM Tris-HCl pH 8.8, 30% v/v glycerol, 2% w/v SDS, 1% w/v DTT) followed by subsequent equilibration in alkylating equilibration buffer (6 M urea, 50 mM Tris-HCl pH 8.8, 30% v/v glycerol, 2% w/v SDS, 2.5% w/v IAA), both 15 minutes at room temperature. The IEF strips were then placed on top of 1 mm 12% SDS-PAGE gels and sealed with agarose. Second dimension SDS-PAGE separation (192 mM glycine, 25 mM Tris, 0.1% w/v SDS) was performed using an Invitrogen X-Cell electrophoresis unit (BioSciences, Dublin, Ireland) using a constant electrophoretic separation power of 1 W/gel. Protein spot features were visualized by Coomassie Brilliant Blue staining. Three replicate gels were run per classification, the excised spot features were pooled from the resulting gels in order to maximize the yield of glycans.

## 3.3 Extraction of IgG from pooled control serum

Serum from healthy volunteers (n = 40, median age = 29 years, 11 males, 29 females), collected at the Clinical Institute Fundeni with informed consent by RNTech (Bucharest, Romania) and distributed as part of the EU FP6 GLYFDIS research consortium, was pooled and used throughout the multiplexed *N*-glycan analysis study, using both UPLC-fluorescence and CE-LIF. IgG was extracted from pooled serum diluted 1:4 with phosphate buffered saline solution, pH 7.4 using an Agilent Technologies Protein A Biomonolith column, (Agilent Technologies, Waldbronn, Germany). After elution with 0.5 M acetic acid the purified IgG was immediately buffer exchanged into 20 mM sodium bicarbonate buffer, pH 7.0 and concentrated *via* centrifugation at 3,000 x g through 5 kDa MWCO filters at 4°C. The purity of the extracted antibodies was assessed using 10% reducing SDS-PAGE with Coomassie Brilliant Blue staining prior to further processing.

## 3.4 Haptoglobin purification for pulmonary disease study

Haptoglobin specific monoclonal antibodies (Biosystems International, Debrecen, Hungary) were purified by Protein-G4 fast flow chromatography (GE-Healthcare, Bucks, UK). The mAb-affinity beads were fabricated by coupling 10 mg of haptoglobin specific monoclonal antibody to 0.5 g

CNBr-activated Sepharose 4B beads according to the manufacturer's instructions (GE-Healthcare). Age and gender matched human control plasma (2 males, average age 61.0), pneumonia plasma (3 males, average age 60.3), COPD plasma (3 males, average age 61.6) and lung cancer plasma (3 males, average age 61.3) were collected with informed consent in the Department of Pulmonology (Debrecen, Hungary) under clinical protocol RKEB /IKEB:2422-2005. The study was approved by the Hungarian Research Ethics Committee. HSA and IgG were depleted using Blue Sepharose 6 and Protein-G4 (GE-Healthcare), respectively. Unbound material was loaded onto the monoclonal antibody affinity column (1 ml bed-volume, in-house made) and incubated for 1 hour at room temperature. The column was washed with 7 column volumes of PBS and haptoglobin was eluted with 2 x 1 ml 0.2 M glycine - HCl (pH 2.7). The eluate was neutralized with the addition of 0.2 ml of 1 M Tris-HCl (pH 9.0). Protein concentration was measured by BCA protein assay kit according to the manufacturer's protocol (Pierce, Rockford, IL, USA). Haptoglobin purity was checked by SDS–PAGE on 4-20% Tris-Glycine precast gradient gels (Invitrogen, Carlsbad, CA, USA) and the concentration was calculated by densitometry analysis using the Gel Logic 2200 PRO equipment and Molecular Imaging software (Invitrogen).

## *3.5 N*-glycan release, labeling and post-derivatization clean-up

*N*-glycans were enzymatically liberated from intact (IgG) or reduced and alkylated glycoproteins (β-glucuronidase and haptoglobin) *via* in solution PNGaseF digestion, the enzyme was used at 10% of the final digest volume (Prozyme). For the stomach cancer study, 5 μL of reduced and alkylated pooled serum in-solution or from 2-DE separated gel spots using in-gel PNGaseF digestion was applied according to the method of Royle *et al.*[225]. Glycans were fluorescently derivatized *via* reductive amination with 2-aminobenzamide (2-AB) and sodium cyanoborohydride in 30% v/v acetic acid in DMSO for UPLC analysis or 8-aminopyrene-1,3,6-trisulphonic acid (APTS) in 15% v/v acetic acid and 1 M sodium cyanoborohydride in tetrahydrofuran for CE-LIF analysis, at 37°C overnight. In both instances the labeling reaction was quenched by the addition of water to yield a final volume of 100 μL. 900 μL of acetonitrile was then added to the reaction vial with subsequent excess fluorophore removal by HILIC phase pipette tips (PhyNexus, San Jose, CA, USA) using an adaptation of the method of Olajos *et al.*[226]. Structural annotation of the chromatographic peaks was performed by comparison of retention time data expressed as glucose units (GU) values with GlycoBase.[86]. *N*-glycan nomenclature and symbolic representations used throughout are as previously described by Harvey *et al.*[2].

## 3.6 Enzymatic *N*-glycan processing and affinity purification

All exoglycosidase digestions were performed following the method of Royle *et al.*[225]. Enzyme digest buffer was 50 mM ammonium acetate (pH 5.5) at 37°C overnight.

In the study associated with phosphorylated oligosaccharides or phosphorylated glycopeptides, selective enriched using the PhosphoCatch™ $TiO_2/ZrO_2$ mixed oxide affinity purification kit (MOAP) was achieved following the manufacturer's instructions (Promega, Madison, WI, USA). In all cases the unretained and the enriched phosphate containing fractions were individually collected and reduced to dryness *via* vacuum centrifugation for subsequent processing and analysis. For qualitative confirmation, an aliquot of the MOAP enriched oligosaccharide fraction and the total *N*-glycan pool were digested with calf intestinal alkaline phosphatase in 50 mM Tris-HCl, pH 9.3, 1 mM MgCl2, 0.1 mM ZnCl2 and 1mM spermidine at 37°C overnight (Promega). Following digestion the enzyme was removed *via* centrifugation through a prewashed 10 kDa MWCO filter. The filtrate was reduced to dryness by vacuum centrifugation prior to analysis.

## 3.7 HILIC UPLC-fluorescence glycan profiling

2-AB labeled *N*-glycans were separated by ultra performance liquid chromatography with fluorescence detection on a Waters Acquity™ UPLC consisting of a binary solvent manager, sample manager and fluorescence detector under the control of Empower 2 chromatography workstation software, (Waters, Milford, MA, USA). Unless otherwise stated, separations were performed using Waters BEH Glycan column, 100 x 2.1 mm i.d., 1.7 μm BEH particles, using a linear gradient of 70-53 % acetonitrile at 0.56 ml/min in 16.5 minutes, 50 mM ammonium formate pH 4.5 was used as buffer A. An injection volume of 20 μL sample prepared in 80% v/v acetonitrile was used throughout. Samples were maintained at 5°C prior to injection and the separation temperature was 40°C. The fluorescence detection wavelengths were λex = 330 nm and λem = 420 nm with a data collection rate of 20 Hz. Retention times were converted into glucose unit (GU) values by time based standardization against a dextran hydrosylate ladder.

## 3.8 Weak anion exchange chromatography based fractionation

Separation of the *N*-glycan pool based upon the degree of sialylation present was performed by weak anion exchange chromatography (AEC) on a GlycoSep C 75 x 7.5 mm, 10 μm DEAE column (Prozyme) using a Waters Alliance 2695 Separations Module with a Waters 474 Fluorescence Detector under the control of Empower Chromatography Workstation (Waters Corporation). A linear

gradient of 100 mM acetate, pH 7.0 in 20% v/v acetonitrile was used for the elution of charged oligosaccharides. Samples for CE-LIF analysis were injected in their unlabeled form and each of the collected fractions was then subsequently labeled with APTS as outlined previously.

## 3.9 Comparative profiling of different linear homo-oligosaccharides

Dextran from *leuconostoc spp.* was hydrolyzed in 0.5 M HCl at 90°C for 1 hour and subsequently subjected to the described HILIC phase purification in pipette tips (PhyNexus). Oligosaccharides were fluorescently derivatized *via* reductive amination with either APTS for CE-LIF investigations, or 2-AB for HPLC studies, as outlined before. CE separations were performed on a Beckman P/ACE MDQ unit (Beckman Coulter, Brea, CA, USA) in a neutral coated N-CHO (PVA) capillary (Beckman Coulter, 50 μm ID / 365 μm OD, 60 cm total, 50 cm effective length), using 25 mM lithium acetate background electrolyte (pH adjusted to 4.75) at 25°C separation temperature (unless stated otherwise) and by applying -20 kV electric potential (E=333 $Vcm^{-1}$). Samples were injected by pressure at 6.9 kPa (1 psi) for 5 seconds. The capillary was pressure rinsed with buffer between injections for 5 minutes at 206.8 kPa (30 psi) in order to avoid any sample carryover. Laser induced fluorescence detection was conducted at $\lambda_{ex}$ = 488 nm and $\lambda_{em}$ = 520. HPLC studies were carried out on a Waters Alliance 2795 Separations Module with a Waters 2475 fluorescence detector under the control of Empower Chromatography Workstation (Waters Corporation). Separations were performed using a TSKgel Amide-80 column (Tosoh Bioscience, Stuttgart, Germany), 250 x 4.6 mm i.d., 5 μm particles, with a linear gradient of 80-42 % acetonitrile at 0.4 mL/min in 160 minutes as previously reported[225]. An injection volume of 20 μL sample prepared in 80 % v/v acetonitrile was used throughout. Fluorescence detection excitation and emission wavelengths were $\lambda_{ex}$ = 330 nm and $\lambda_{em}$ = 420 nm, respectively. Structural models of malto- and cello-oligosaccharides, energy minimized by the application of the GLYCAM06 force field parameters[227], were obtained *via* the Carbohydrate Builder (GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA, USA) and visualized in Discovery Studio version 2.5 (Accelrys, San Diego, CA, USA).

## 3.10 CE-LIF *N*-glycan profiling

Capillary electrophoresis separations of APTS labeled *N*-glycans were performed on a Beckman PA800 Plus Pharmaceutical Analysis System under the control of 32 Karat software, version 9.0 (Beckman Coulter) using a N-CHO neutral coated capillary, 360 μm O.D., 50 μm I.D. Unless otherwise stated the total capillary length was 60 cm with a 50 cm effective separation length from the injection

to the detector. The instrument was equipped with a 3 mW 488 nm solid state laser and a 520 nm cut off filter for laser induced fluorescence detection. For all experiments a commercially available carbohydrate separation buffer was used (Beckman Coulter) at a separation temperature of 25°C and using an applied potential of -30 kV, (500 $Vcm^{-1}$), reversed polarity. Prior to use, the capillary was conditioned by subsequent rinsing with water and carbohydrate separation buffer for 10 minutes at 30 psi. In order to avoid any possible carryover, the capillary was also rinsed with buffer for 5 minutes at 30 psi between injections. To improve precision, two bracketing standards were used, APTS labeled maltose as the lower and 2-aminoacridone labeled glucuronic acid (2-AMAC GA) as the upper bracket. Samples were injected by pressure at 1 psi for 5 sec. Migration times were converted to relative migration times using the bracketing standard boundaries. These normalized migration times were subsequently converted to GU values using a fifth order time based standardization against a malto-oligosaccharide ladder (Grain Processing Corporation, Muscatine, IA, USA). In order to automate this process a script was written in Matlab version 7.6 (The Math Works, Natick, MA, USA).

## 3.11 Protein ID with LC-MS/MS for stomach cancer study

Deglycosylated 2-DE gel spots were washed repeatedly with acetonitrile and 50 mM ammonium bicarbonate buffer and digested overnight with trypsin (Promega modified sequencing grade, MSC, Dublin, Ireland). Peptides were eluted from the gel pieces by successive stepwise washing with an increasing concentration of acetonitrile and reduced to dryness *via* vacuum centrifugation. LC-MS/MS analysis was performed using an Agilent Technologies 1200 series nano scale HPLC instrument consisting of nano and capillary binary gradient pumps and a micro well plate autosampler connected *via* the Agilent Technologies ChipCube interface to an Agilent Technologies 6340 series ion trap mass spectrometer operated in positive ion mode with a spray voltage of -1.8kV (Agilent Technologies). Separations were performed using a ProtID-Chip-43 consisting of a 40 nL enrichment column and a 43 mm analytical column packed with Zorbax StableBond 300 Å 5 μm C18 particles. Peptides were eluted using a linear gradient of 5-60 % acetonitrile containing 0.1 % v/v formic acid in 6 minutes following a 1 minute isocratic hold at 5 % acetonitrile. The mass spectrometer was operated in automatic data dependant switching mode with active exclusion. The three most abundant MS ions were selected for MS/MS analysis. MS/MS data was analyzed using Agilent Technologies Spectrum Mill Proteomics Workbench version A.03.03.084 SR4 (Agilent Technologies, Santa Clara, CA, USA) and searched against the National Centre for Biotechnology Information database using the human mouse taxonomic filter. Search parameters used included a precursor ion mass tolerance of 2.5 Da, product ion mass tolerance of 0.7 Da, cysteine carbamidomethylation specified as a fixed

modification, asparagine deamidation and methionine oxidation as variable modifications and a maximum of two missed cleavage sites allowed. Returned protein identifications were automatically validated using the auto validation feature of Spectrum Mill, peptide scores > 6, protein scores > 11 and the percentage scored peak intensity > 60%.

## *3.12 N*-Glycosylation site analysis of β-glucuronidase

*Protease digestion for N-glycosite analysis*. 100 μg aliquots of β-glucuronidase were reduced, alkylated and digested with either Asp-N in 100 mM ammonium bicarbonate buffer or Lys-C in 25 mM Tris HCl, pH 8.5, 1 mM EDTA overnight at 37°C (Roche Diagnostics, Mannheim, Germany). Digests were reduced to dryness *via* vacuum centrifugation, reconstituted in 50 μL of 0.1% v/v trifluoroacetic acid (TFA) and the resulting peptides separated, collected and processed to characterize the *N*-glycosylation site to which oligosaccharides displaying mannose-6-phosphate were attached.

*LC-UV peptide mapping*. Peptide separations were performed using a Waters Alliance 2695 Separations Module with a Waters 2489 UV/Visible detector under the control of the Empower chromatography workstation software (Waters Corporation). A Waters XBridge 3.5 μm, 135 Å $C_{18}$ 150 x 2.1 mm i.d. column was used for all separations. Peptides were eluted using a linear gradient of 5-65% acetonitrile containing 0.1% v/v TFA in 55 minutes at a flow rate of 200 μL/min. The column temperature was 40°C and detection was performed at 214 nm. All resulting chromatographic peaks were individually collected, deglycosylated and subjected to glycosylation profiling by UPLC and peptide identification using LC-MS/MS.

*Peptide identification using LC-MS/MS*. LC-MS/MS analysis of enriched deglycosylated peptides was performed using a Dionex Ultimate 3000LC (Sunnyvale, CA, USA) connected to a LTQ Orbitrap XL mass spectrometer (ThermoFisher Scientific, Hemel Hempstead, UK). Separations were performed using an increasing linear gradient of acetonitrile on a Dionex $C_{18}$ Pepmap, 0.075 x 150 mm column at 300 nL/min. The mass spectrometer was operated in the positive ion mode, the spray voltage was -1.8kV, and the transfer capillary temperature 200°C. Data was acquired in the automatic data dependent switching mode. A high resolution accurate mass MS scan was performed using the Orbitrap to select the 5 most intense ions in the 300-2000 m/z range prior to MS/MS analysis in the linear ion trap. Resulting MS/MS data were searched using BioWorks 3.2 (ThermoFisher Scientifc) against the Uniprot-Swissprot database with the Homo sapiens taxonomic filter specified. The enzymatic deamidation of asparagine was manually annotated in a separate β-glucuronidase sequence FASTA file added to the database. Peptides were filtered according to the following parameters: XCorr ≥1.90 (+1), 2.0 (+2) and 2.5 (+3). Identified spectra were also verified manually.

## 3.13 Statistical analysis

Non parametric Kruskall-Wallis analysis of all glycomic data was performed using SPSS statistical analysis software Version 15.0 for Windows to evaluate differences in the median values across pathological groups with each experimental parameter. Differences between variables were further investigated using with Mann-Whitney or Students t-tests. In all instances a P value $< 0.05$ was considered as being statistically significant.

# CHAPTER 4

# Results

## 4.1 UPLC serum *N*-glycome profiling to identify cancer associated changes

### 4.1.1 Separation of the serum *N*-glycome using HILIC-UPLC with fluorescence detection

Here, a newly available sub-2 μm HILIC phase for UPLC was evaluated for the separation of serum *N*-glycans. Figure 6 displays an overlay chromatogram of fluorescently labeled *N*-glycans released from control serum and separated using different column lengths run under chromatographic conditions scaled according to the column length. Feature numbers, *i.e.* the number of distinct and integratable peaks using the integration parameters as outlined previously in each chromatogram, were determined to be 30, 46 and 53, at 10, 20 and 30 minutes analytical run times (including column re-equilibration), for the 5, 10 and 15 cm analytical columns, respectively.

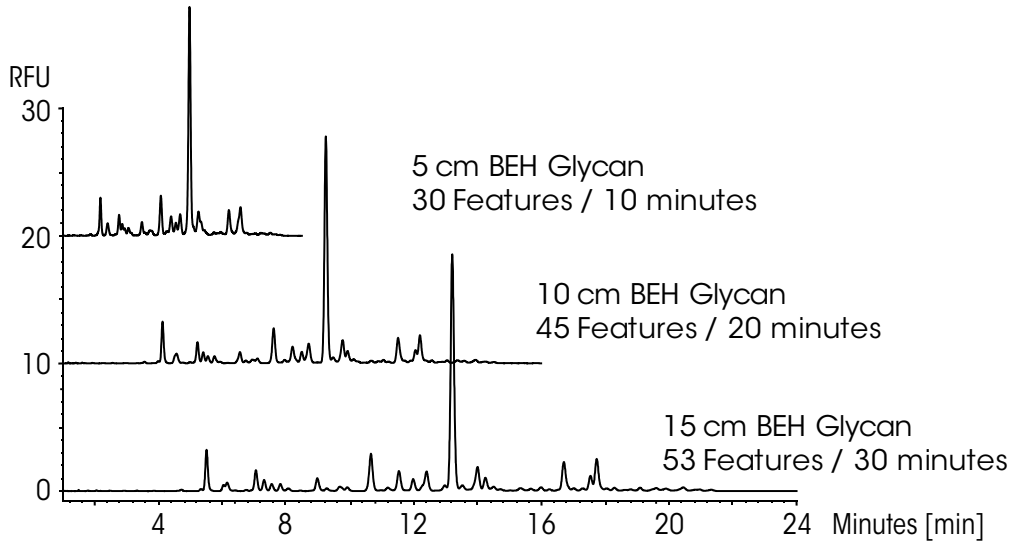



***Figure 6:*** *Evaluation of the selectivity of different column lengths of the 1.7 μm BEH Glycan UPLC HILIC column. Adapted from*[69].

The chromatographic efficiency of the sub-2 μm HILIC phase permitted the separation of positional and linkage specific isomeric oligosaccharides and also glycans containing a bisecting *N*-acetyl glucosamine

residue. Figure 7 displays an expansion of the regions on either side of the largest chromatographic peak along with two highlighted critical peak pairs. Using the sub-2 µm HILIC phase it was possible to separate the core fucosylated biantennary galactosylated glycan FA2G2 from the core fucosylated biantennary galactosylated analogue containing the bisecting *N*-acetyl glucosamine residue FA2BG2 and partially separate out the disialylated core fucosylated biantennary galactosylated glycan FA2G2S(6)2 and its analogue containing a bisecting *N*-acetyl glucosamine residue.

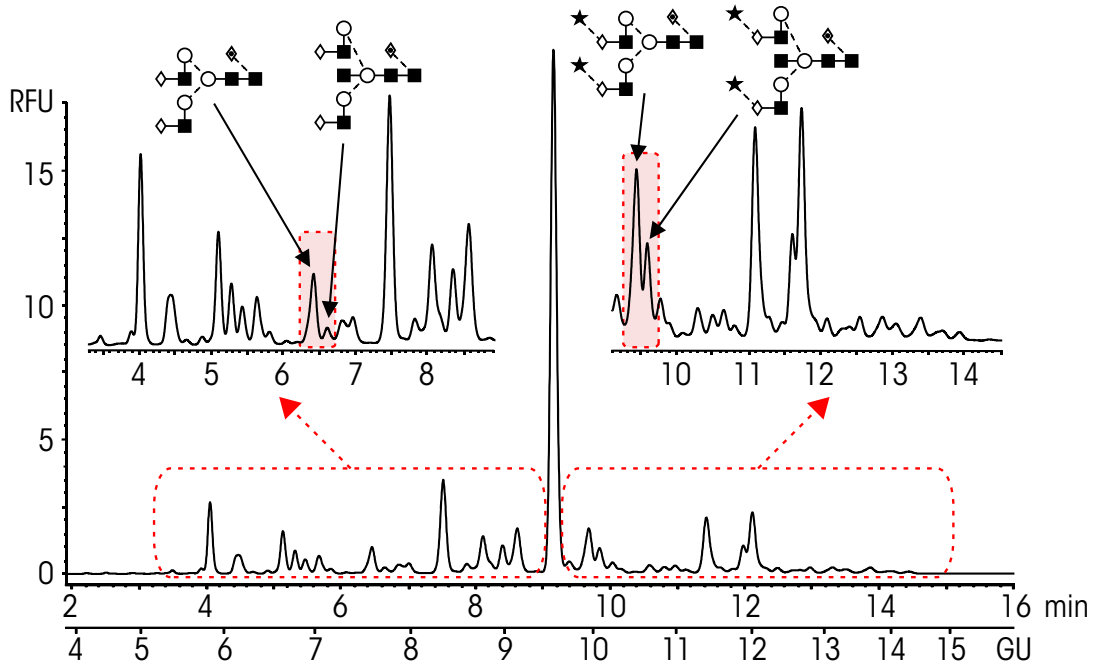



***Figure 7:*** *Separation of the serum N-glycome on the 100 × 2.1 mm BEH Glycan column with expansions of the regions of the chromatogram on either side of the main peak (A2G2S2). Highlighted in each expanded region is the separation of a critical glycan pair demonstrating the superior efficiency of BEH Glycan and its ability to resolve analogous oligosaccharides differing only in the presence of a bisecting N-acetyl glucosamine residue. Adapted from*[69].

### 4.1.2 UPLC-fluorescence rapidly identifies cancer associated alterations in glycosylation

During the initial column evaluation the 15 cm BEH Glycan column was noted to operate at the higher end of the pressure limits. For all subsequent experiments the 10 cm BEH Glycan column was used. Serum *N*-glycans from the pooled pathological staged patient groups were injected in triplicate onto the 10 cm BEH glycan column, the resulting data was aligned and integrated. In total, 35 out of the 45 *N*-glycan peaks were returned as being statistically altered *versus* pathological classification. To evaluate the trends further, box-plots were also generated. Figure 8 displays a selection of the resulting box-plots and highlights the *N*-glycans which displayed the most distinct trends in expression.

Of note from Figure 8 is the reduction in the levels of asialo- and monosialylated core fucosylated *N*-glycans with disease progression. The levels of disialylated afucosylated biantennary *N*-glycans increased in an almost linear fashion with increasing disease progression. For the core fucosylated disialylated biantennary *N*-glycan a 'U' shaped trend in the box-plot was observed demonstrating an initial decrease in the levels of this *N*-glycan in the early stages of carcinogenesis with a subsequent increase as the disease progresses. Similar increases in the levels of trisialylated triantennary glycans were also observed. Also of note is the tight distribution within each box-plot which is an indication of high method precision.

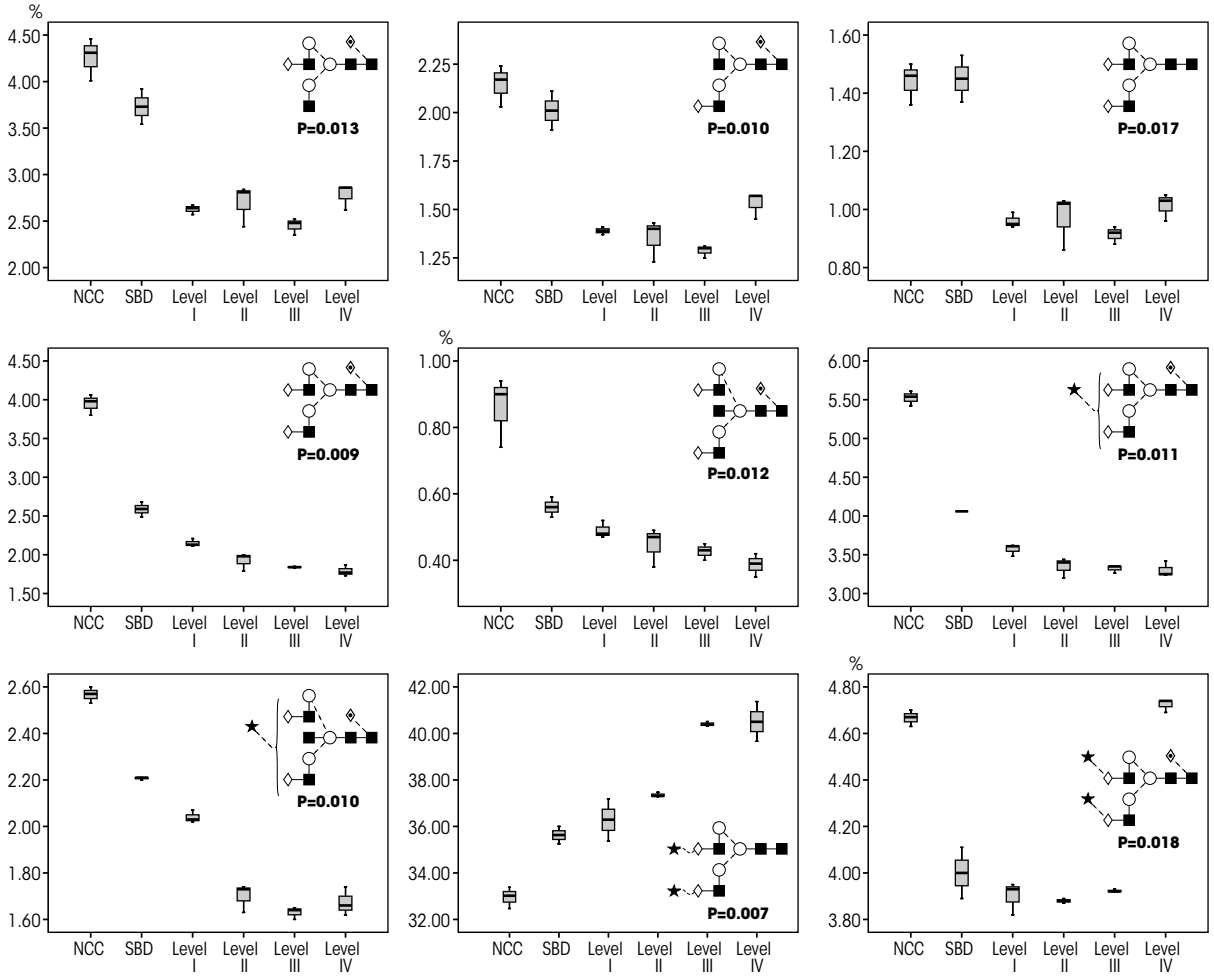



***Figure 8:*** *Box-plots indicating the trends in expression of a number of glycans as indicated with disease progression. Abbreviations: NCC refers to the noncancer patient cohort; SBD, patients with benign stomach disease; and Level I to Level IV stomach cancer AJCC, American Joint Committee on Cancer staging system (sixth edition). Adapted from[69].*

### 4.1.3 Contribution of abundant glycoproteins to the serum *N*-glycome

The contribution of four high abundant (>1 mg·mL$^{-1}$ concentration) glycoproteins to the serum *N*-glycome, *i.e.* IgG, haptoglobin, transferrin and α1-acid glycoprotein, was investigated. *N*-glycans were released from commercially available IgG, haptoglobin, transferrin and α1-acid glycoprotein extracted

from human serum or plasma (Abcam, Cambridge, UK or Sigma-Aldrich, Dublin, Ireland). The released *N*-glycans were then profiled on the 10 cm BEH Glycan column and compared against the *N*-glycan profile of control serum. Figure 9 displays the resulting overlaid chromatograms and the contribution of the glycans from the individual glycoproteins to the serum *N*-glycome trace. IgG appears to contribute the majority of asialo glycans and glycans bearing a core fucose to the serum trace. *N*-glycans from transferrin contribute mono- and disialo- biantennary glycans. Glycans present on haptoglobin are predominantly biantennary disialylated and triantennary trisialylated whilst α1-acid glycoprotein contributes biantennary disialylated, triantennary trisialylated and tetraantennary tetra sialylated glycans to the serum *N*-glycome. Additionally, control serum was depleted of these and other abundant proteins using the Agilent Technologies MARS Human-14 immunoaffinity depletion column. As depicted in the bottom trace of Figure 9, the depletion process removes the majority of core fucosylated asialo *N*-glycans and results in a reduction in the levels of tetra antennary tetrasialo *N*-glycans. The most abundant glycan present both before and after immunoaffinity depletion of abundant proteins remains the disialo biantennary afucosylated glycan. The levels of triantennary trisialo *N*-glycans remain constant after depletion.

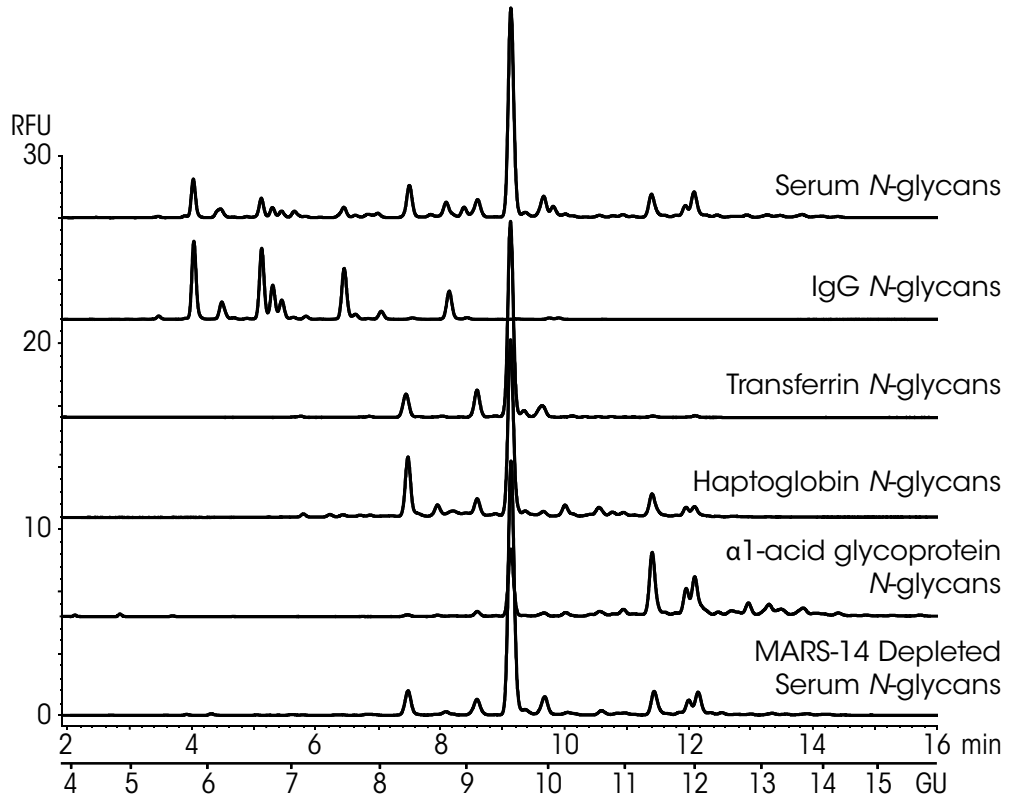



***Figure 9:*** *Evaluation of the contribution of glycosylation present on four abundant glycoproteins to the total serum N-glycome in addition to the effect of the depletion of these glycoproteins from the serum using a commercially available multiple residue immunoaffinity column. Adapted from*[69].

### 4.1.4 Monitoring altered glycosylation with cancer progression using HILIC-UPLC-fluorescence.

Having investigated the contribution of four highly abundant glycoproteins to the serum glycome it was decided to evaluate whether the glycosylation present on these glycoproteins displayed any potential clinical utility to monitor carcinogenesis and progression of gastric adenocarcinoma. IgG was extracted from the pooled patient serum using Protein A (Pierce Protein A Spin Plates, Rockford, IL, USA). For haptoglobin, transferrin and α1-acid glycoprotein isolation, 150 µg of serum proteins were separated using small scale two dimensional electrophoresis (2-DE). Following protein spot visualization *via* Coomassie Brilliant Blue staining, the spot features corresponding to haptoglobin, transferrin and α1-acid glycoprotein were excised. Figure 10 displays a reference image of the 2-DE separation of the non cancer patient serum proteins. The identity of the excised proteins was verified by proteomic analysis using LC-MS/MS, the resulting data is presented in Table 1. The first dimension IEF range chosen was pH 4-7 in order to maximise the resolution of both the haptoglobin and transferrin spot features and to improve separation from other closely migrating proteins. As a result of this, it is acknowledged that the more acidic isoforms of α1-acid glycoprotein, pI 2.8-3.8[228] may have potentially been lost during the IEF focusing step. Overlaid chromatograms as generated using the sub-2 µm HILIC phase for the glycans released from each glycoprotein from the chosen pathologically staged patient pools are displayed in Figure 11.

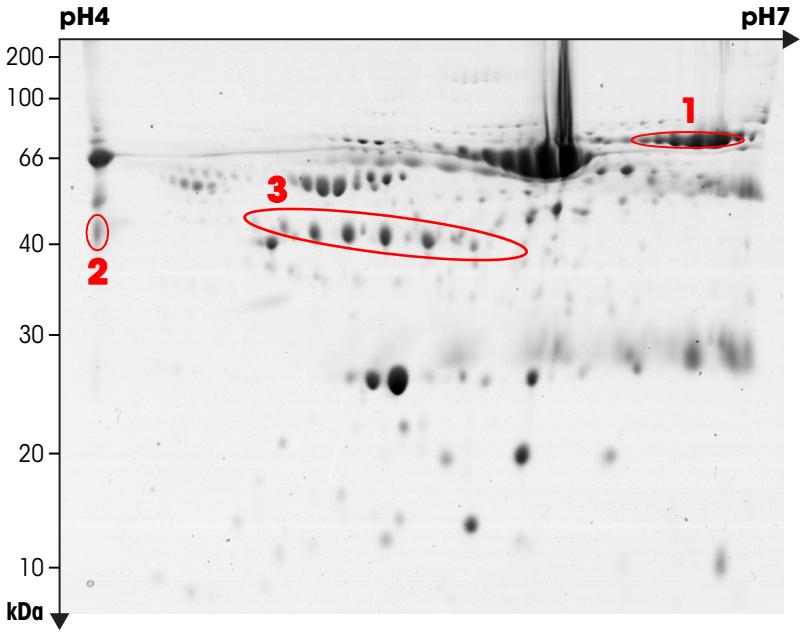



***Figure 10:*** *Separation of 150 µg of total serum proteins present in control serum using two-dimensional electrophoresis, IEF pH 4-7 followed by 12% SDS-PAGE, visualization using Coomassie Brilliant blue staining. The spot features as highlighted were excised for glycomic evaluation in all pathologically staged gels. LC-MS/MS based proteomic identification data for the proteins present is listed in Table 1.*

***Table 1:*** *LC-MS/MS based proteomic identification of the excised 2-DE spots as indicated on Figure 10.*

| Spot Cut | Protein Identified | Uniprot Accession Number | Summed MS/MS Score | % Coverage | No. of Distinct Peptides |
|---|---|---|---|---|---|
| 1 | Transferrin<br>IGHM Protein | P02787<br>P01871 | 424.85<br>49.82 | 43<br>7 | 24<br>3 |
| 2 | α1-AGP<br>Haptoglobin<br>Complement C3 | P02763<br>P00738<br>P01024 | 130.8<br>86.21<br>32.88 | 41<br>12<br>1 | 8<br>5<br>2 |
| 3 | Haptoglobin<br>Apolipoprotein A-IV | P00738<br>P06727 | 153.94<br>29.42 | 23<br>7 | 8<br>2 |

Examination of the chromatographic data revealed a number of alterations present in the *N*-glycan profiles of each of the glycoproteins purified from the pathologically staged serum pools. For the extracted IgG the most noticeable alterations in glycosylation, as depicted in Figure 11A, included an increase in the levels of the core fucosylated agalactosyl *N*-glycan (FA2) with disease progression accompanied by a decrease in both the asialo- mono- and bigalactosylated core fucosylated *N*-glycans (FA2[6]G1, FA2[3]G1, FA2G2). For haptoglobin, transferrin and α1-acid glycoprotein the most noticeable alterations in glycosylation as depicted in Figure 11B, C and D respectively, were an increase in the levels of sialylated glycans with disease progression. For haptoglobin the levels of both peaks corresponding to the sialic acid linkage isomers of the disialylated triantennary glycan were decreased with tumor growth with an increase in the levels of the trisialylatyed triantennary glycan noted. Increases in fucosylation were also observed, higher levels of the monosialo biantennary core fucosylated glycan (FA2G2S1) were present on transferrin in both cancer stages than in the benign conditions. Also for haptoglobin and α1-acid glycoprotein an increase in the peak containing a triantennary trisialylated *N*-glycan bearing outer arm fucosylation in the form of a sialyl Lewis X ($SLe^X$) epitope was observed to increase with advancing pathogenesis.

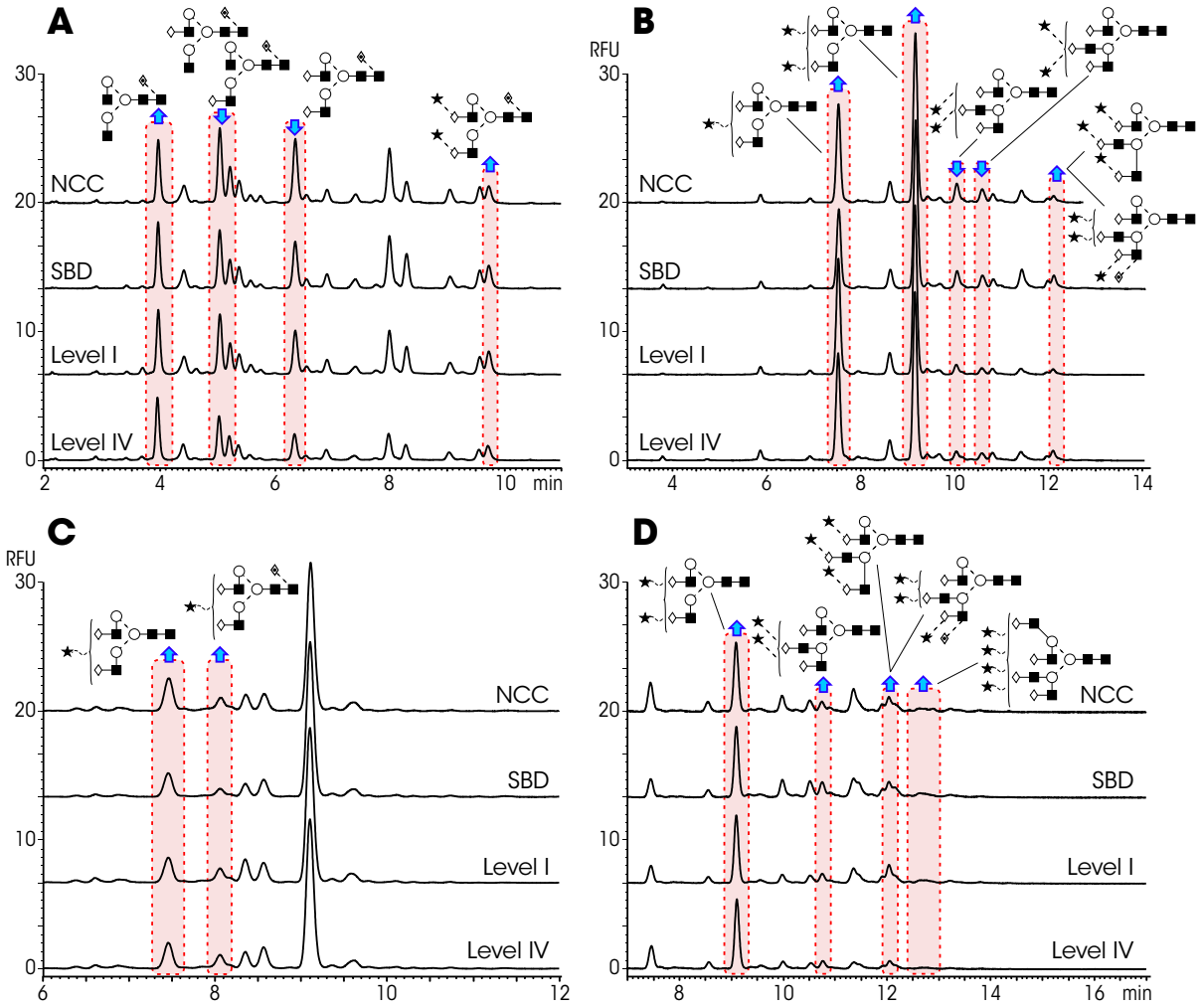



***Figure 11:*** *Alterations in the N-glycosylation present on (A) IgG, (B) haptoglobin, (C) transferrin, and (D) α1-acid glycoprotein as purified from the pathologically staged serum as indicated using either affinity purification or two-dimensional electrophoresis. Abbreviations: NCC refers to the noncancer patient cohort; SBD, patients with benign stomach disease; Level I and Level IV stomach cancer AJCC, American Joint Committee on Cancer staging system (sixth edition). Adapted from[69].*

## 4.2 Robust and reproducible profiling of isomeric oligosaccharides by CE

As opposed to surface interaction based methods, such as HILIC, capillary electrophoresis (CE) represents an electric field mediated separation technique, offering different selectivity. Moreover, since CE separation is based on charge to hydrodynamic radius ratios, it is capable of separating positional and linkage isomers with great efficiency, especially when compared to chromatographic methods which are based on surface interaction principles. Therefore, differences in linkage and/or antennary position and associated molecular conformation of oligosaccharides play a significant role in their electrophoretic separation.

### 4.2.1 Experimental resolution and precision

The resolving power of CE is demonstrated in Figure 12 comparing the separation of three closely related oligosaccharide structures of glucose based ladders differing only in their linkage type (α1→4, β1→4 and α1→6). While the migration of monomers (open ring conformation glucose due to the conjugated derivatization agent at the reductive end in each instance) was not influenced by the different linkage types, the separation behavior of the corresponding species with higher degrees of polymerization (DP≥2) diverged. Migration differences between alpha- and beta-linked oligosaccharides with the same DP were more pronounced, whereas isomeric α1→4 and α1→6 linked species exhibited almost identical migration behavior in the lower degree of polymerization range of DP 2-6.

The alignment of the three different homooligosaccharides based upon free APTS as lower and APTS labeled maltoheptaose (G7) as upper bracketing standard is also depicted in Figure 12. Average %RSD for migration times (n=3 repetitions) of malto- (DP 1-25), cello- (DP 1-6) and isomalto- (DP 1-25) oligosaccharides were 0.02, 0.08 and 0.05, respectively; RMT deviations within the bracketing boundaries (DP 1-6) were 0.02, 0.03 and 0.04 %, respectively.

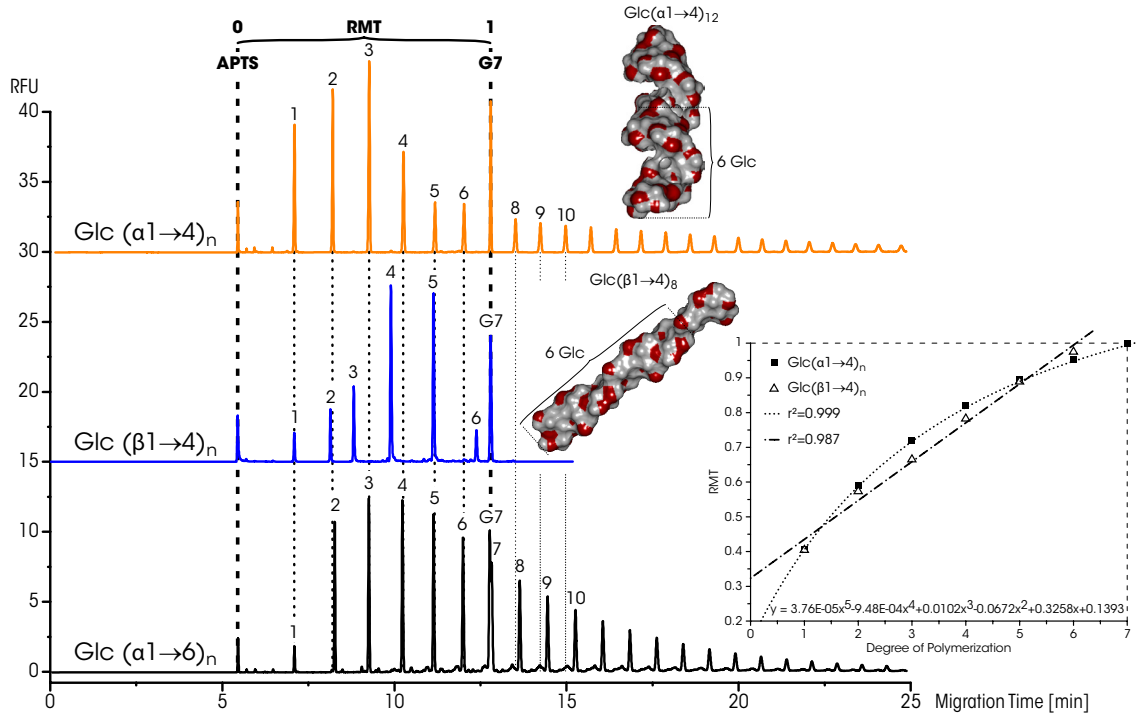



***Figure 12:*** *CE-LIF separations of APTS labeled malto- Glc(α1→4)$_n$, cello- Glc(β1→4)$_n$ and isomalto-oligosaccharide Glc(α1→6)$_n$ ladders. Numbers above the peaks represent the degree of polymerization of the species in each ladder. Bracketing boundaries, APTS as lower and APTS labeled maltoheptaose (G7) as upper marker are indicated with the curly bracket and bold dashed lines. Energy minimized models of malto- (DP 12) and cello- (DP 8) oligosaccharides are also shown. Inset shows a 5$^{th}$ order polynomial fitting for Glc(α1→4)$_n$ with the corresponding equation and a linear fitting for Glc(β1→4)$_n$. Adapted from*[229].

### 4.2.2 Glucose unit values

Table 2 provides the calculated sugar unit values for cello- Glc(β1→4)$_n$ and isomaltooligosaccharide Glc(α1→6)$_n$ ladder peaks based on their relative migration to the maltooligosaccharide Glc(α1→4)$_n$ ladder, using both eq. 4 (Theoretical Background section) and a 5$^{th}$ order polynomial fitting equation (latter results shown in brackets), referred to as glucose units (GU). Although, glucose was the building block in all three ladders, the inherent different linkage types of the three homooligomers rendered the corresponding GU values to the individual isomeric oligomers to be different, thus reflected their electromigration differences.

Accurate fitting of the Glc(β1→4)$_n$ homooligosaccharide ladder (DP 1-6, $r^2$=0.987) could be obtained by simply mapping a line onto the data points as shown in the inset in Figure 12, thus no higher order polynomial approximation was needed. The attempt of a linear fitting of Glc(α1→4)$_n$ (DP 1-6) on the other hand resulted in a rather poor correlation ($r^2$=0.935) and therefore a 5$^{th}$ order polynomial approach was used ($r^2$=0.999).

***Table 2:*** *Calculated glucose unit (GU, based on the maltooligosaccharide ladder) values for cello- and isomalto-oligosaccharide samples using eq. 4 (Theoretical Background section) and a 5$^{th}$ order polynomial fitting equation (shown in brackets).*

| DP | Glc(α1-4)$_n$ | Glc(β1-4)$_n$ | Glc(α1-6)$_n$ |
|---|---|---|---|
| 1 | 1.000 (1.000) | 0.984 (0.987) | 0.994 (0.995) |
| 2 | 2.000 (2.001) | 1.907 (1.891) | 2.020 (2.018) |
| 3 | 3.000 (2.998) | 2.573 (2.534) | 2.970 (2.963) |
| 4 | 4.000 (4.003) | 3.624 (3.593) | 3.956 (3.953) |
| 5 | 5.000 (4.997) | 4.929 (4.917) | 4.951 (4.942) |
| 6 | 6.000 (6.002) | 6.464 (6.444) | 5.971 (5.970) |

### 4.2.3 Molecular conformation impact on electromigration

The changes in electromigration of Glc(α1→4)$_n$ and Glc(α1→6)$_n$ oligosaccharides with an increasing degree of polymerization are compared in Figure 13. The resulting degree of polymerization *versus* migration time plot exhibited a curved increase for both oligosaccharide ladders in the low degree of polymerization range. Maltooligosaccharides from 1-7 sugar units showed an average increase of 1 min per ΔGU, changing to a linear increase of ~ 0.70 min per ΔGU after around 7-8 glucose units. This transition to a linear-like migration time increment per ΔGU is indicated by the excellent least-squares fit from GU 9 throughout 29 ($r^2$=0.9995). Isomaltooligosaccharides exhibited almost identical migration behavior as

their α1→4 linked counterparts in the 1 to 6 DP range. Although, commencing with DP 7, the migration behavior of malto- and isomaltosides diverged. The inset in Figure 13 shows the differential migration times of adjacent oligo-sugar units, thus further reflects the altering relation between degree of polymerization and hydrodynamic volume equivalent radius. Differential migration curves split after DP 6 where α1→6 oligosaccharides started to exhibit linear-like distances between adjacent oligomers. Unit-to-unit distances of maltooligosaccharides on the other hand, further decreased at DP>6 probably due to the lower hydrodynamic radius of the α1→4 linked species.

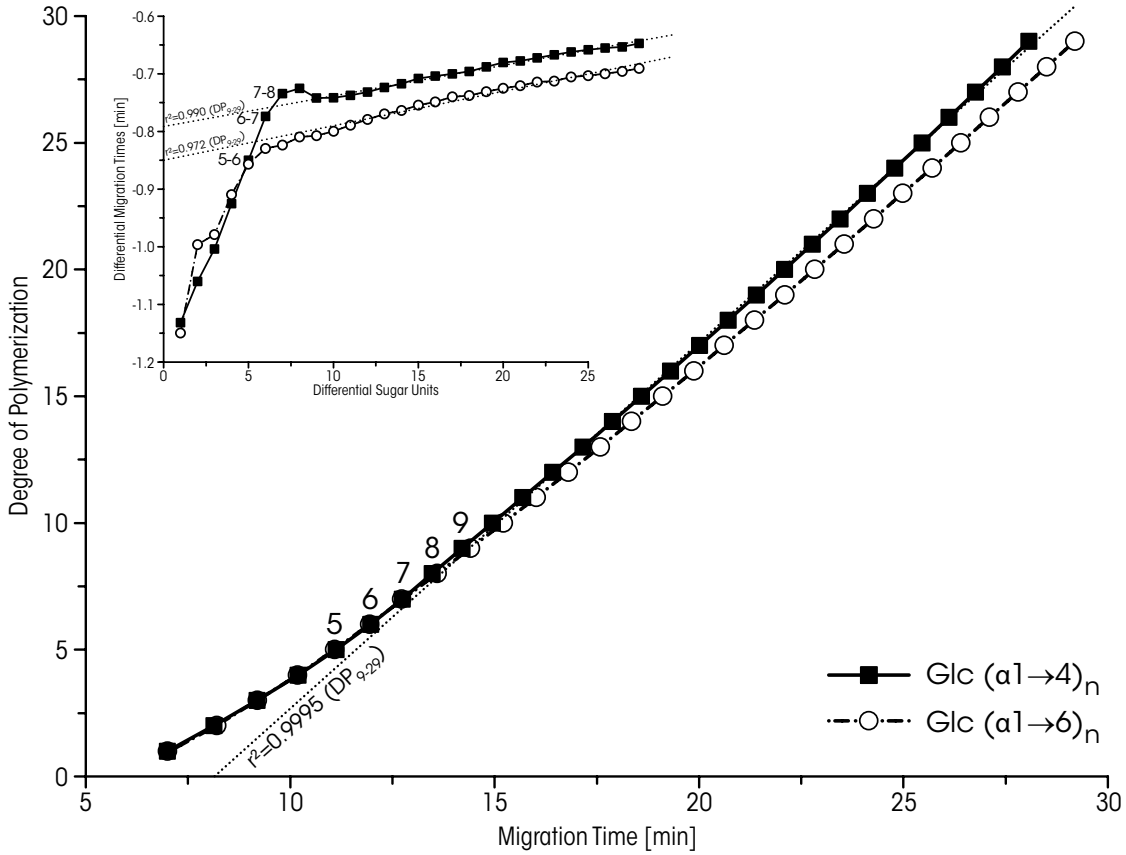



***Figure 13:*** *Electromigration changes of Glc(α1→4)$_n$ and Glc(α1→6)$_n$ homooligomers with increasing degree of polymerization. Inset shows the differential migration times of adjacent oligo-sugar units. Numbers indicate degree of polymerization, and the dotted lines represent least-square fittings. Adapted from*[229].

### 4.2.4 From electromigration theory to geometry

Numerous attempts (mainly for peptides) have been undertaken to simplify the fundamental electrophoretic mobility equation given in eq. 2 (Theoretical Background section), with the goal of a deeper understanding of the relationship between the electromigration behavior and molecular properties, for spherical rigid solutes, associating the electrophoretic mobility ($\mu$) to their molecular mass ($M$)[230]:

$$\mu = q \cdot M^{-x} \qquad \text{(eq.5)}$$

Based on eq. 5, the electrophoretic mobility is directly proportional to the charge ($q$) of the solute and

inversely proportional to its molecular mass. Exponents ($x$) in most instances were reportedly ranging between 1/2 and 2/3. The exponent 2/3 suggests a random coil configuration, considering the real shape of the coil in solution with electrostatic and other interactions, respectively; whereas conformations of linear chains exhibit an electrophoretic mobility that is inversely proportional to $M$, *i.e.* x=1[231].

To further investigate the molecular shape differences of the sugar homooligomers manifested in their electromigration, the slope values of logarithmic electrophoretic mobility (log($\mu$)) *versus* logarithmic molecular mass (log($M$)) plots, presented in Figure 14, were evaluated. Linear fittings in two separate sections of lesser and greater than DP 9, resulted in high correlations $r^2$=0.999 and 0.999, respectively (Panel A). For comparative reasons, the same intervals were chosen for isomaltooligosaccharides (Panel B). Also beta-linked oligosaccharide ladders, Glc(β1→4)$_n$ and GlcNAc(β1→4)$_n$, exhibited a bend between the data points corresponding to DP 3 and 4 (Panels C and D, respectively). Table 3 provides the resulted slope values of the linear fitting lines between the specified DP ranges for each oligosaccharide ladder.

***Table 3:*** *Comparison of slope values from log(μ) versus log(M) plots of homooligosaccharide ladders in the specified degree of polymerization range, at 25°C separation temperature.*

| Slope Values | | DP | $r^2$ |
|---|---|---|---|
| Glc(α1→4)$_n$ | -0.62 | 1-9 | 0.999 |
| | -0.68 | 9-24 | 0.999 |
| Glc(α1→6)$_n$ | -0.63 | 1-9 | 0.999 |
| | -0.71 | 9-24 | 0.999 |
| Glc(β1→4)$_n$ | -0.51 | 1-3 | 0.987 |
| | -0.81 | 3-6 | 0.997 |
| GlcNAc(β1→4)$_n$ | -0.51 | 1-3 | 0.982 |
| | -0.81 | 3-6 | 0.999 |

Maltosides with α1→4 type linkages exhibited a slope value of -0.66 between DP 9-17 (data not shown) and -0.68 in the higher degree of polymerization range (DP 9-24). Isomaltosides with α1→6 linkages had comparable slope values in the DP 1-9 range as the corresponding maltosides (-0.63), but exhibited a lower slope value of -0.71 in the higher DP range (DP 9-24). Unfortunately, only low molecular size β-linked homooligosaccharide ladders were available for this study (DP 1 to 6), thus allowed only limited comparison. Although, β1→4-linked oligo-glucoses in the very low DP range (DP 1-3) were also considered to be affected by the labeling reaction (slope: -0.51), higher degree of polymerization structures (DP 3-6) exhibited slope values of -0.81. Also the sterically larger *N*-acetyl glucosamine (GlcNAc(β1→4)$_n$) ladder (due to the acetyl functionalities present at $C_2$) exhibited identical slope values to its glucose counterpart as shown in Table 3.

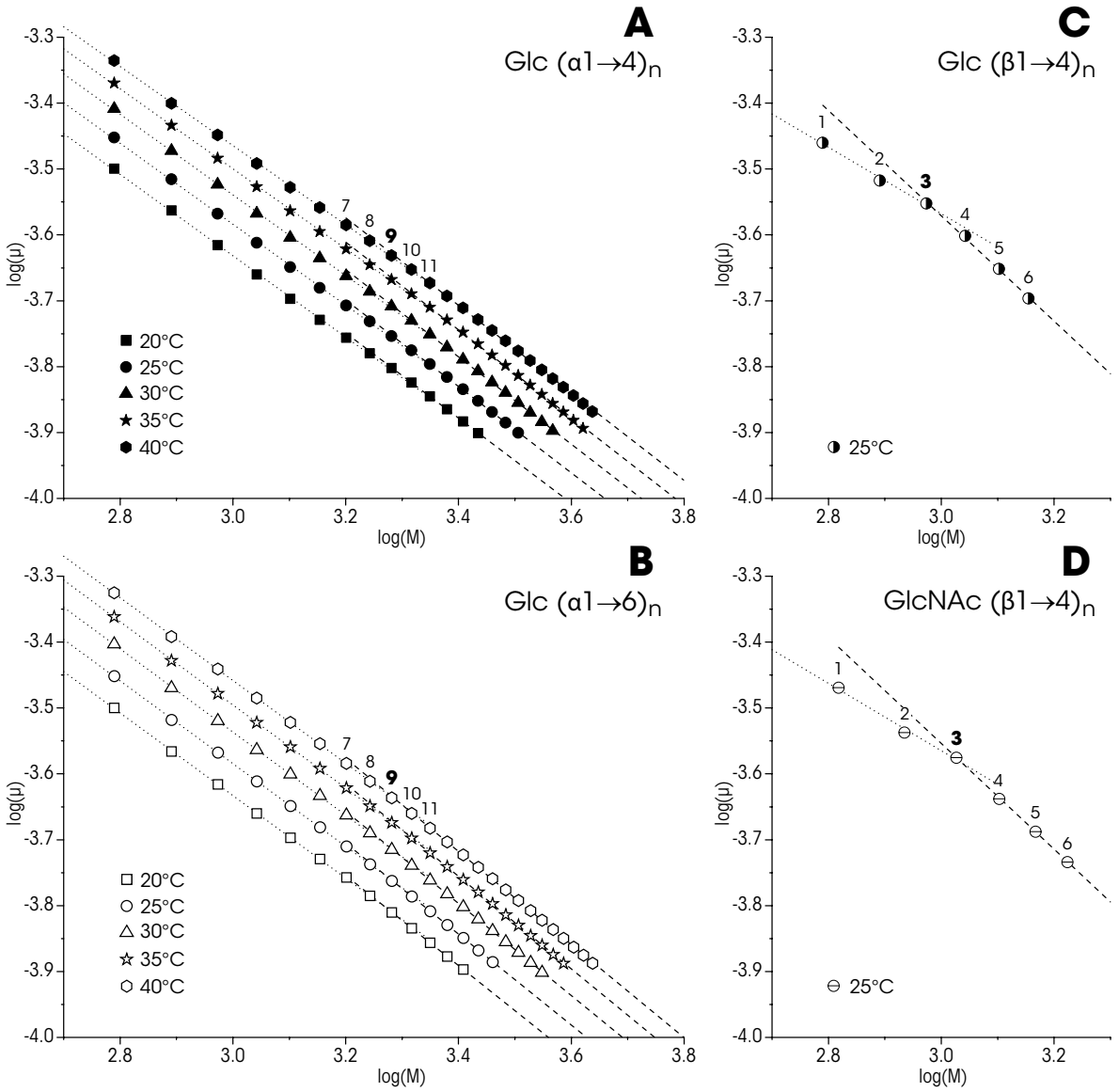



***Figure 14:*** *Log(μ) versus log(M) plots of malto- (A) and isomalto- (B) oligosaccharides, both measured at 20, 25, 30, 35 and 40°C separation temperatures. Panels C and D show the same plots for cello- and N-acetylchitooligosaccharides at 25°C. Numbers represent degree of polymerization, dotted and dashed lines are linear fittings before and after the bend indicated in bold (A and B = 9; C and D = 3). Adapted from*[229].

### 4.2.5 Temperature effects on electromigration and associated conformation

To investigate the impact of temperature on the electrophoretic separation of oligosaccharides, both malto- and isomaltosides were examined at 20, 25, 30, 35 and 40°C separation temperatures, respectively. Due to the limited molecular size range availability (DP 1 to 6), β-linked homooligosaccharide ladders were not subjected to the temperature study. Electrophoretic mobility *versus* temperature showed linear increases for every oligosaccharide throughout the investigated DP range of 1-14 (data not shown). In addition, the comparison of the slope values of log($\mu$) *versus* log($M$) plots, as

depicted in Figure 14, did not reveal remarkable alterations in the two intervals below and above DP 9 ($r^2 \geq 0.999$ in all instances). Slope values of α1→4 linked oligo-glucoses slightly increased in the DP 1-9 range presenting -0.617 at 20°C, -0.615 at 25°C, -0.611 at 30°C, -0.606 at 35°C and -0.600 at 40°C, whereas α1→6 isomaltosides exhibited a slope value of -0.63 throughout. In the higher degree of polymerization ranges, malooligosaccharides in contrast presented a descending tendency of their slope values of -0.645 at 20°C (DP 9-14), -0.655 at 25°C (DP 9-17), -0.662 at 30°C (DP 9-19), -0.665 at 35°C (DP 9-23) and -0.664 at 40°C (DP 9-24). Isomaltooligosaccharide on the other hand slightly decreased in slope values of -0.683 at 20°C (DP 9-14), -0.690 at 25°C (DP 9-16), -0.699 at 30°C (DP 9-20), -0.703 at 35°C (DP 9-22) and 0.707 at 40°C (DP 9-26).

### 4.2.6 Presentation and shielding of hydroxyl groups

With a view to evaluate the relative differences of presenting hydroxyl groups which are not shielded by hydrogen bridges, malto- and isomaltohomooligosaccharide ladders were subjected to hydrophilic interaction liquid chromatography analysis. The resulting chromatograms are depicted in Figure 15. In contrast to the CE data, the oligo-glucoses with inherent α1→4 and α1→6 glycosidic linkages exhibited differential retention starting as low as from DP 2.

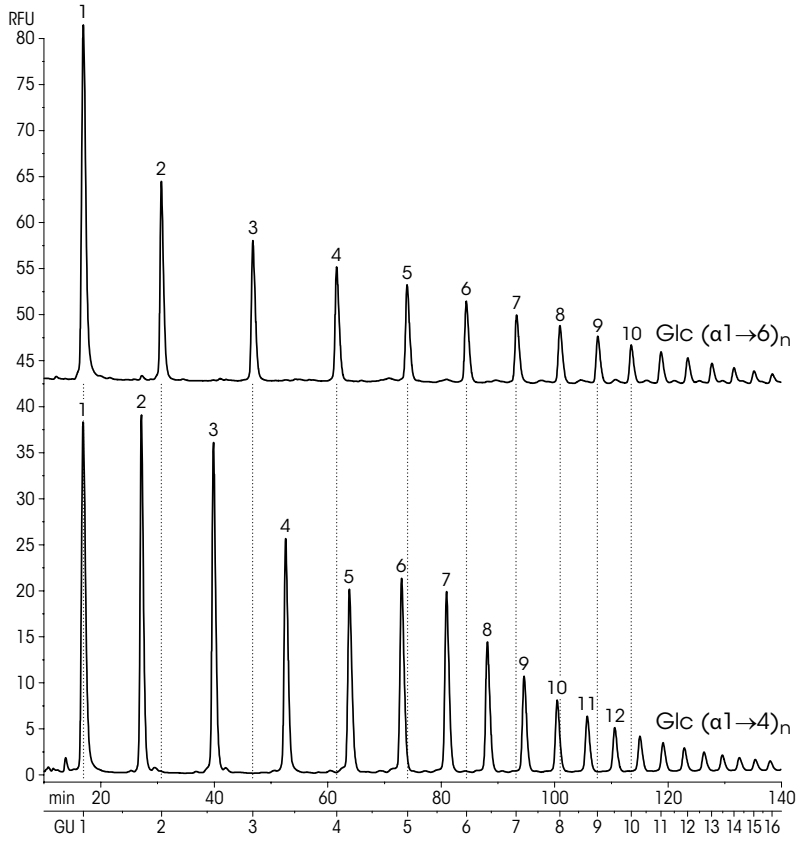



***Figure 15:*** *HILIC separations of 2-AB labeled isomalto- Glc(α1→6)$_n$ and malto-oligosaccharide Glc(α1→4)$_n$ ladders (A). Adapted from*[229].

An increased retention of isomaltosides compared to maltosides in the HILIC column can be observed from the chromatograms in Figure 15 and the alterations in retention with increasing degree of polymerization shown in Figure 16A. Changes in the curvature of DP *versus* retention (Figure 16A) and differential retention times (Figure 16B), were observed in the DP range between 6 and 8. The remarkable negative slope between the first and the second data point in Figure 16B indicated the differences between a 2-AB labeled single open ring sugar molecule and its higher degrees of polymerization.

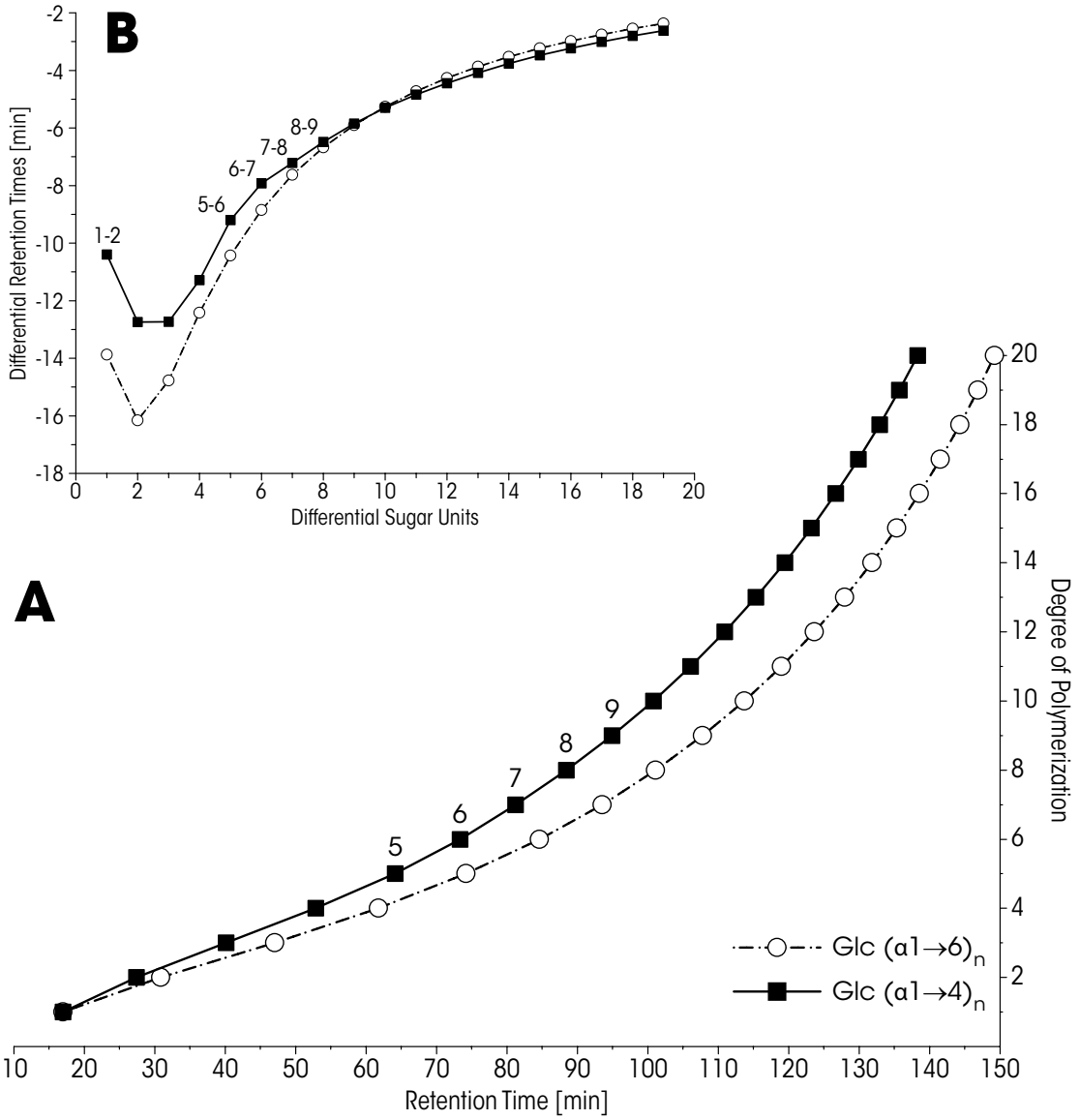



***Figure 16:*** *HILIC retention time changes of homooligosaccharide ladders with increasing degree of polymerization (A). Differential retention times of adjacent oligo-sugar units (B). Numbers indicate the degree of polymerization. Adapted from*[229].

## 4.3 CE profiling of haptoglobin *N*-glycan alterations in pulmonary diseases

The previously developed (section 4.2 and[229]) robust method for carbohydrate analysis by capillary electrophoresis with laser induced fluorescence detection (CE-LIF) is applied. Differences in the *N*-glycosylation patterns of a small scale cohort of acute- and chronic-inflammatory as well as malignant pulmonary disease patients are investigated. After albumin and IgG depletion, haptoglobin was partitioned from the plasma samples of patients by affinity purification. The purity of the partitioned haptoglobin was checked by SDS-PAGE. *N*-glycans were then released by *N*-glycanase, fluorescently labeled with APTS and analyzed by CE-LIF after the removal of the excess dye.

### 4.3.1 Exoglycosidase digestion-based structural annotation of haptoglobin *N*-glycans

The excellent resolving power and high detection sensitivity of CE-LIF offers the ability to separate both positional and linkage carbohydrate isomers, based on their hydrodynamic volume to charge ratio differences, as outlined before (section 4.2 and[229]). The highly complex *N*-glycan profile of a commercially available haptoglobin protein standard was first analyzed and compared to the maltooligosaccharide ladder as depicted in Figure 17, traces A and B. A total of 27 peaks were observed in trace B, including eight low abundant species with < 2% relative peak areas. Removal of sialic acid residues by sialidase digestion resulted in a significant shift of the observed peaks towards the longer migration time zones of the electropherogram, apparently due to the loss of the extra negative charges sialic acids held, as shown in Figure 17, trace C. The benefit of this step was twofold, (i) simplification of the electropherogram; and (ii) more importantly, increase in sensitivity as problems with disease related analyte dilution over multiple peaks due to the presence of sialic acid linkage isomers was alleviated. Alterations in sialylation have been considered to be disease associated; however, the high separation efficiency of CE greatly complicated the annotation and interpretation of the resulting data. While efforts are ongoing to address this analytical complexity, at present the best solution was to remove the sialic acids and focus on the presence of other disease associated alterations. The nine peaks observed after sialidase digestion (Trace C) were identified as follows: fully galactosylated bi- (A2G2), tri- (A3G3) and tetra- (A4G4) antennary glycans with additional core fucosylation (FA2G2, FA3G3, FA4G4) or antennary / arm-fucosylation (A2FG2, A3FG3, A4FG4). Structural elucidation was accomplished by co-injection (spiking) with purified *N*-glycan standards, sequential exoglycosidase digestion and database analysis of the associated shifts expressed in glucose units (GU). GU values were calculated for all nine asialo haptoglobin glycan peaks by applying a fifth order time based standardization against the malto-oligosaccharide ladder (Trace A). CE-LIF profiles from sequential exoglycosidase array digestion with α1-3 (arm) specific fucosidase, α1-6 (core) specific

fucosidase, β1-4 galactosidase and β1-2,3,4,6 hexosaminidase are compared in traces D - G of Figure 17, respectively. After the treatment with the mixture of all above listed exoglycosidases, the digestion profile of complex type *N*-glycans was expected to comprise of only the pauci-mannose (MAN3) structure (Trace G). By the theoretical reattachment of the previously digested monosaccharide residues, the basic antennary-structures (A2, A3, A4) in trace F, fully galactosylated species (A2G2, A3G3, A4G4) in trace E, corresponding glycans containing antennary-fucosylation in trace D and core-fucosylation in trace C, were all identified. These structures of interest were subsequently recognized in the clinical samples by horizontal alignment based on the internal standard and comparison of the corresponding GU values.

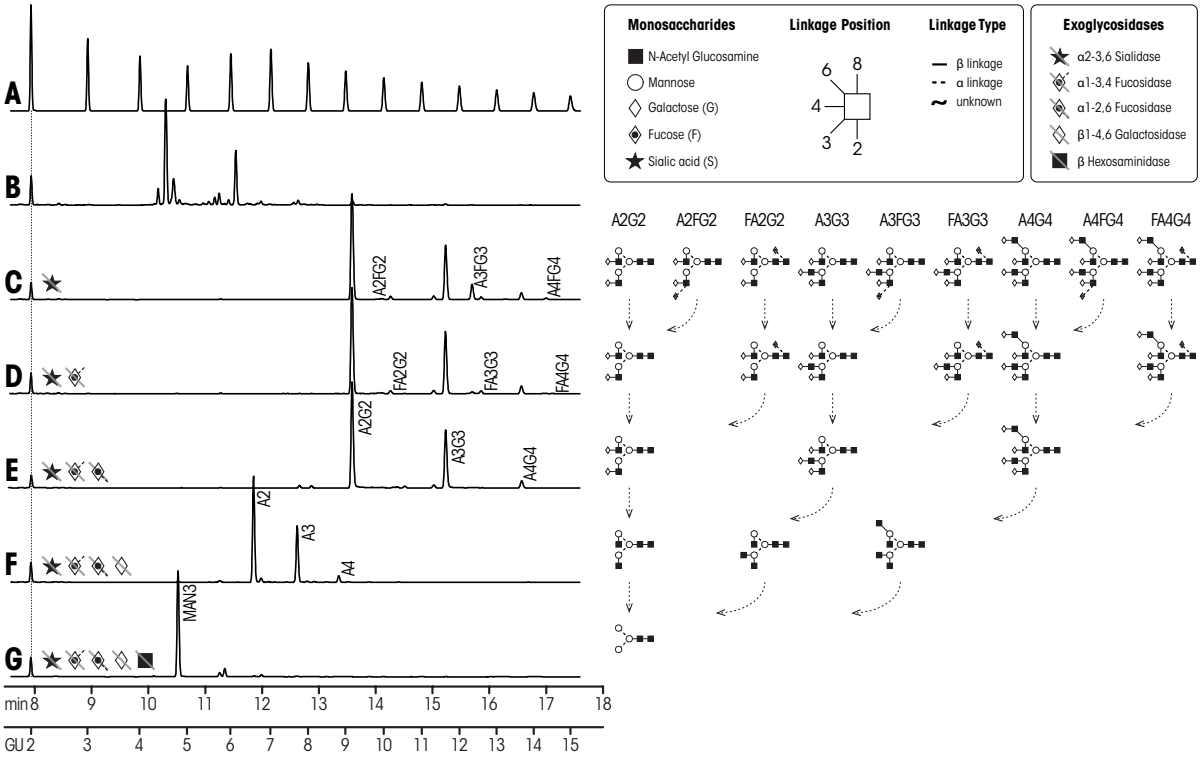



***Figure 17:*** *Identification of the N-linked glycans of standard haptoglobin using exoglycosidase array digestion and CE-LIF analysis. Trace A: Malto-oligosaccharide ladder, Trace B: Undigested Haptoglobin N-glycan profile, Sequential digestion by α2-3,6 sialidase (Trace C), α1-3/4 fucosidase (Trace D), α1-6 fucosidase (Trace E), galactosidase (Trace F) and hexosaminidase (Trace G). Adapted from*[232].

### 4.3.2 Isolation of haptoglobin from human plasma samples

Haptoglobin was isolated from the plasma samples of 2 control, 3 pneumonia, 3 COPD and 3 age and gender matched lung cancer patients by means of haptoglobin specific antibody affinity chromatography as specified in the experimental section. The high concentration of serum albumin and IgG in the samples necessitated their depletion before the affinity chromatography step. Approximately

50 µg haptoglobin was isolated from each patient plasma sample with 55-90% purity, checked by SDS-PAGE. The *N*-glycans of the captured haptoglobins were released by PNGase F digestion after purification, then APTS labeled and analyzed by CE-LIF.

### 4.3.3 Analysis of haptoglobin *N*-glycans in patient samples

Using the relative percentage area values of the identified glycan structures, disease associated alterations were determined in the control, acute and chronic inflammation as well as cancer patient samples by statistical analysis. The box-plots in Figure 18 present the trends in expression of a number of glycans between the indicated patient groups. For instance, the comparison of control *versus* COPD, presented in Figure 18B and C, revealed a significant decrease in the amount of A2G2 glycan and an increase in A3G3 glycan, also indicating elevated branching in this group.

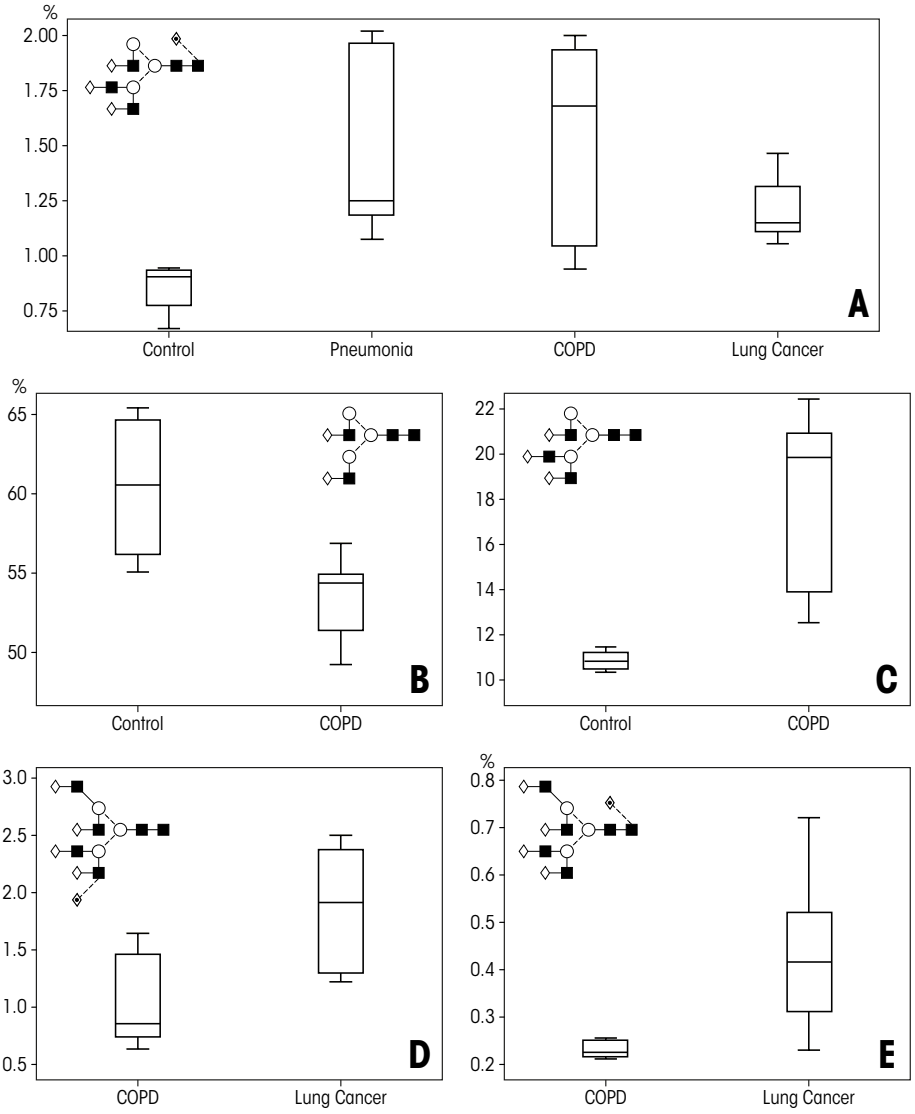



***Figure 18:*** *Box-plot comparison exhibiting the trends in expression of a number of glycans between the indicated groups. Adapted from*[232].

The two types of fucosylation on human haptoglobin are core-fucosylation, where the fucose unit is linked (α1-6) to the reducing end GlcNAc residue that is directly connected to the polypeptide backbone, and antennary-fucosylation where the fucose unit is linked to any of the antennary GlcNAc residues *via* α1-3 or 4 linkages. The relative amounts of the different fucosylation type containing structures (*i.e.*, antennary and core) were also compared between the different patient groups. As shown in Figure 18A, the elevation of core fucosylated, galactosylated tri-antennary (FA3G3) glycan level was clearly discriminative between the control and disease groups. Statistical analysis was also carried out among the disease groups, revealing interesting differences between COPD and lung cancer. A noteworthy increase was detected in the amount of core (FA4G4) and antennary (A4FG4) fucosylated tetra-antennary glycans in lung cancer compared to COPD with the respective p values of 0.009 and 0.02 as shown in Figure 18D and E.

## 4.4 Selectivity of UPLC *vs.* CE-LIF: Structure elucidation of IgG *N*-glycans

### 4.4.1 Experimental design and performance

IgG was extracted from healthy human control serum using Protein A and *N*-glycans were liberated *via* PNGase F digestion. Isolated glycans were fluorescently labeled and profiled using UPLC-fluorescence and CE-LIF after excess label removal. An overview of the analytical approach used in the current study is shown in Figure 19. IgG from human serum was specifically chosen and the Fc were on purpose not separated from the Fab regions due to the potentially higher degree of oligosaccharide structural complexity compared to many therapeutic monoclonal antibodies such as those produced in unmodified CHO cells.

To maximize the resolving power of both UPLC-fluorescence and CE-LIF column and capillary lengths were optimized, respectively. For the 1.7 μm HILIC phase an increase in the column length resulted in an increase in efficiency as indicated in Figure 20 (A) generated using scaled gradient conditions. For rapid profiling or screening, *e.g.* during clone selection, where full structural characterization is not necessary, the *N*-glycan profile generated using the 5 cm column is considered to be sufficient. The reduced resolution between the core fucosylated (F) bi-antennary (A2) monogalactosylated (G1) positional isomers, FA2[6]G1 and FA2[3]G1, more commonly referred to as G1 and G1' and their analogues bearing a bisecting GlcNAc, complicated individual structural quantitation. Optimum separation performance was recorded using the 10 cm column. Increasing the column length from 10 to 15 cm resulted in only a minor increase in efficiency accompanied by a longer runtime and increased backpressure; 800 bar on the 15 cm column as compared with 575 bar on the 10 cm column at a mobile phase composition of 53% acetonitrile and a flow rate of 0.56 mL/min.

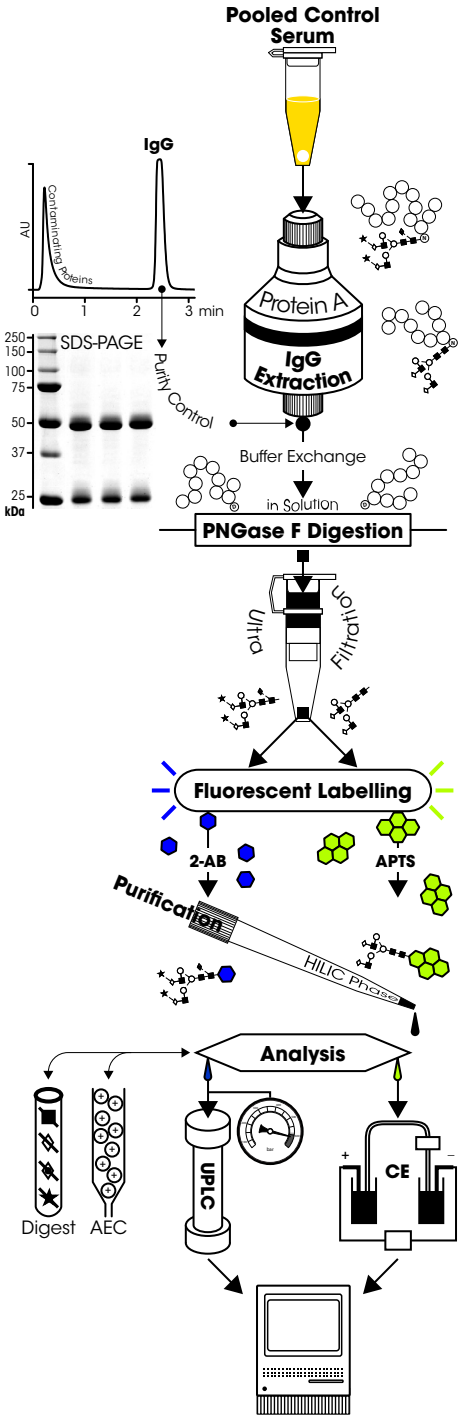



***Figure 19:*** *Schematic of the sample preparation and analysis routes of section 4.4. Adapted from*[233].

In contrast, CE-LIF demonstrated superior efficiency due to baseline separation of the core fucosylated monogalactosylated positional isomers using both a 40 and a 50 cm effective capillary length as shown in Figure 20 (B). The increase of capillary length from 40 to 50 cm showed a shoulder on the peak corresponding to FA2[3]G1. Further inspection using exoglycosidase digestion revealed that the shoulder was due to co-migration with two other oligosaccharides, the structurally related

analogue containing a bisecting GlcNAc residue, FA2B[6]G1 and the afucosyl bigalactosylated glycan, A2G2. UPLC-fluorescence displays enhanced selectivity over CE-LIF in separating these oligosaccharides, wherein it was possible to separate all these structurally related glycans as indicated in the region between 5 and 6 minutes in the main trace of Figure 20 (A). Both HILIC and CE-LIF required the same injection-to-injection runtime but CE-LIF required a shorter separation window than UPLC-fluorescence.

To further exemplify the annotation of the critical positional isomer peak pair and A2G2, tailored exoglycosidase digests were performed using (i) hexosaminidase and (ii) α1-2/3 mannosidase[158]. The digests were profiled using both UPLC-fluorescence and CE-LIF. Electropherograms for the identification of these positional isomers are inserted under the main trace in Figure 20 (B). The dashed lines indicate alignment to the lower and upper bracketing standards. Using these sequential digests, the individual structures in the critical peak pair could be identified, as FA2[6]G1 upon hexosaminidase treatment generates a substrate for the α1-2/3 mannosidase that results in a peak shift to the left as indicated, whereas FA2[3]G1 does not.

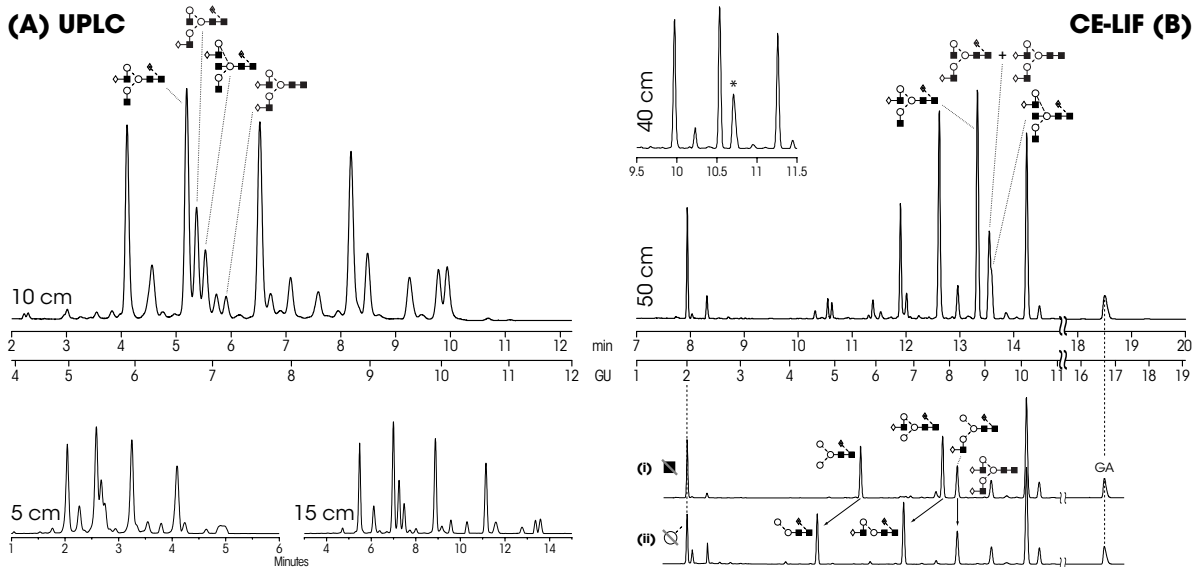



***Figure 20:*** *Comparison of (A) UPLC-fluorescence using various column lengths and (B) CE-LIF using a 50 cm effective / 60 cm total capillary length for the separation of N-glycans released from polyclonal serum IgG. Inset in (B) is the separation of the same IgG N-glycan pool using a 40 cm effective / 50 cm total capillary length. Underlaid in (B) is the identification of the G1 positional isomers using sequential digestion with (i) β-1-2/3/4/6 hexosaminidase and (ii) α1-2/3 mannosidase with peak identification and annotation. Adapted from[233].*

### 4.4.2 Structural annotation using exoglycosidase digestion

Exoglycosidase array digestions were performed to identify the peaks present in the UPLC-fluorescence and CE-LIF IgG *N*-glycan profiles. Subsequent profiling of the digestion products was conducted using both separation techniques. After desialylation the resulting CE-LIF profiles exhibited similar selectivity to UPLC-fluorescence, which facilitated structural annotation. Figure 21 displays the resulting electropherograms accompanied by the experimentally annotated oligosaccharide digestion pathway for the IgG *N*-glycan pool. Each electropherogram was aligned using the lower and upper bracketing standards, however, for clarity only the migration window between ten and fifteen minutes is displayed. The structural annotation of the peaks corresponding to APTS labeled asialo glycans was consistent with that generated using UPLC-fluorescence. For these APTS labeled asialo glycans, structural annotation was performed using a bottom up approach considering the knowledge that the peaks present in the lowermost trace corresponded to the bi-antennary *N*-glycan and its analogue carrying a bisecting GlcNAc residue.

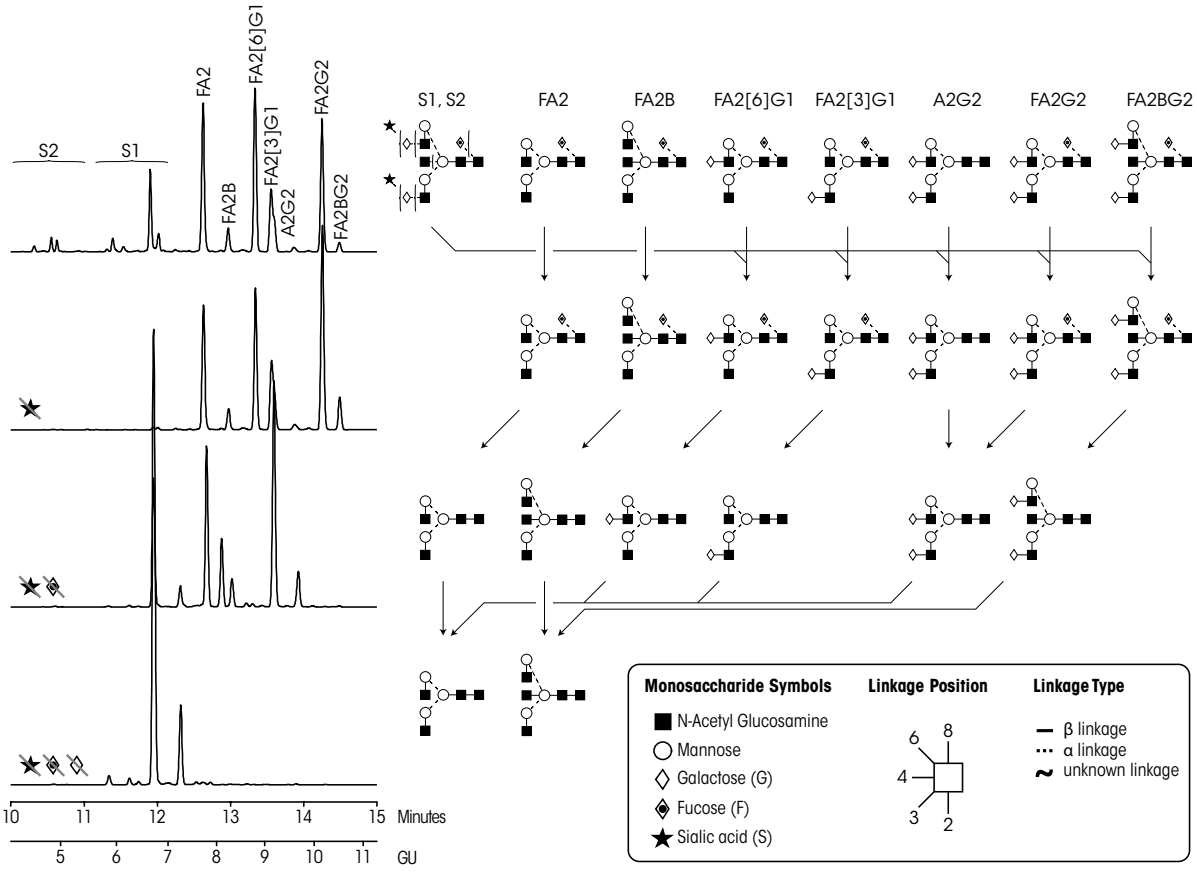



***Figure 21:*** *The experimentally annotated exoglycosidase digestion pathway of the IgG N-glycan pool using CE-LIF for the profiling of the resulting digest products. Adapted from*[233].

The GU shifts following the consecutive removal of terminal sugar units when using UPLC-fluorescence were in agreement to those previously reported, *i.e.* removal of a β1-4 linked galactose residue caused a shift of ~0.8 GU and removal of an α1-6 linked fucose or a β-linked GlcNAc residue resulted in a shift of ~0.5 GU[225]. A noteworthy observation in CE-LIF was that the GU shifts following digestion showed deviations based upon the presence of other structural components due to the resulting differences in the hydrodynamic volume. The removal of an α1-6 linked fucose from the core fucosylated biantennary bigalactosylated glycan, FA2G2, resulted in a GU shift of 1.01, whereas the removal from its structural analogue containing a bisecting GlcNAc residue resulted in a GU shift of only 0.86. Previous *in silico* based studies on the molecular dynamics and conformation of such glycans have revealed that removal of the fucose residue results in differential spacing of the antennae in the bisecting GlcNAc containing A2BG2 glycan compared to A2G2, due to a decrease in the distance between the antennary galactose residues following defucosylation[234-235]. The galactose residues were also closer together in the absence of the bisecting GlcNAc. The information suggested by molecular modeling can be experimentally observed in the current data with discrete alterations in the GU shifts based upon the differences in the molecular volumes of the underlying oligosaccharide structures. Retention on the HILIC phase is dependent upon the number of hydrophilic groups presented[106]. As the bisecting GlcNAc residue is contained within the cleft of the two antennae, it is only partially available for interaction resulting in a small increase in retention time / GU value.

A further advantageous feature of CE-LIF was minimal sample consumption per injection. The injection volume is on average three orders of magnitude lower than that onto the UPLC column. For the characterization of recombinant antibodies, sample quantities are generally not limited; however, for the profiling of the glycosylation present on antibodies extracted from low volume clinical specimens, the reduced consumption of CE-LIF is an attractive feature. Accompanied with this is the ability to generate additional information output from the same sample using CE-LIF due to the possibility of sequential digest and re-injection.

### 4.4.3 Sialoform identification by hyphenation with offline charge fractionation

The presence of sialic acids has orthogonal effects on the separation of glycans in UPLC and CE-LIF. The addition of sialic acids results in increased molecular charge and subsequently faster electromigration, whereas in HILIC retention increases with increasing glycan size and associated hydrophillicity. Digestion with sialidase and successive monitoring of the peak shifts to the asialo counterpart positions can be difficult if not impossible for lower abundant sialoforms due to only minor increases in the respective relative peak area in the asialo forms.

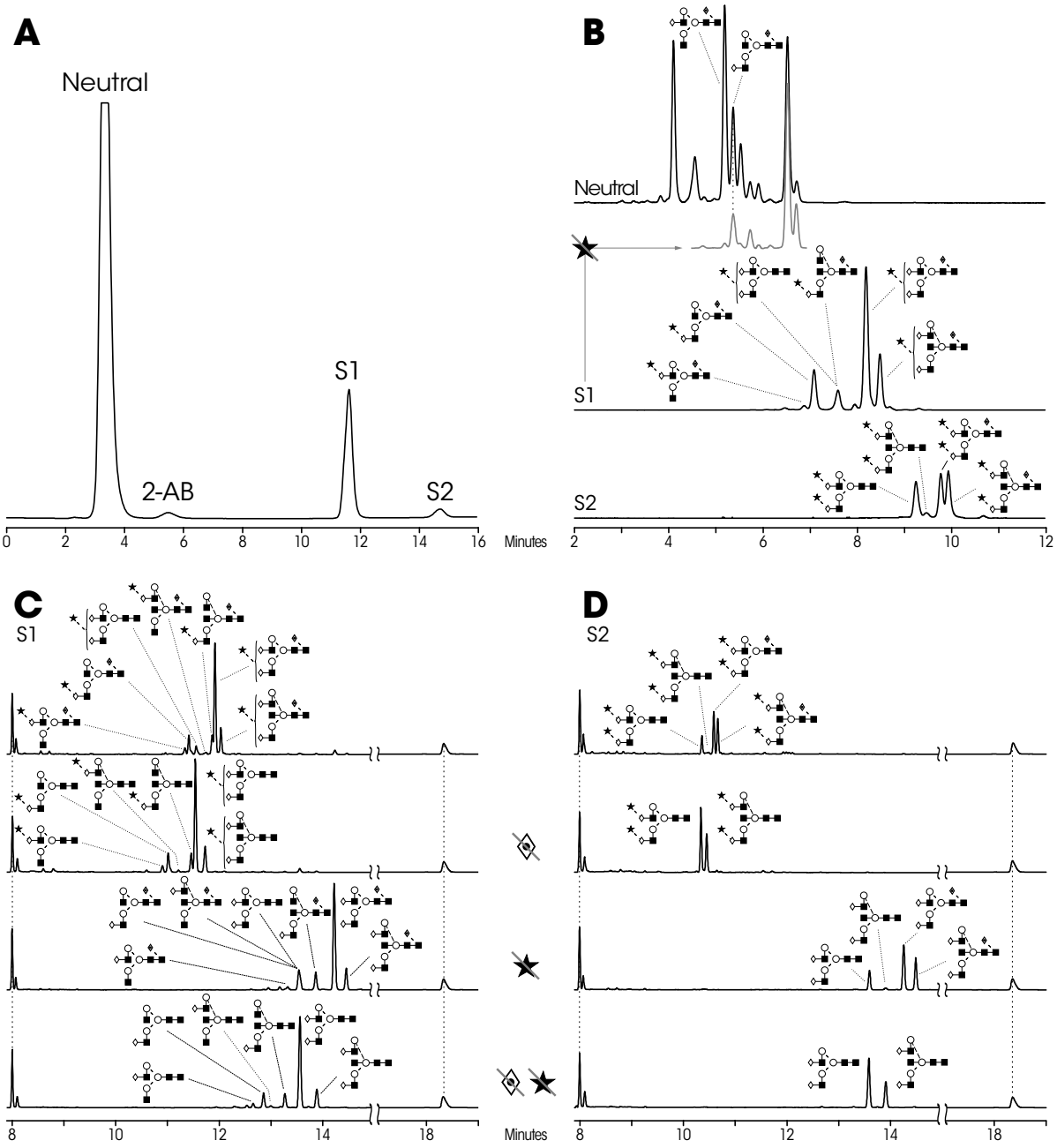



***Figure 22:*** *(A) Fractionation of the IgG N-glycan pool using weak anion exchange chromatography. (B) Profiling of the resulting fractions using UPLC-fluorescence with annotation of the resulting peaks. Inset is the α2-3/6/8 sialidase digest of the S1 fraction. (C) Profiling of the S1 fraction by CE-LIF along with digestion individually with α1-2/3/4/6 fucosidase and α2-3/6/8 sialidase and subsequently combined α1-2/3/4/6 fucosidase and α2-3/6/8 sialidase along with the annotation of the resulting peaks. (D) Profiling of the S2 fraction by CE-LIF along with the same digestions as were performed for the S1 fraction. Adapted from*[233].

Exhaustive structural annotation of the peaks corresponding to sialylated oligosaccharides in both the UPLC-fluorescence chromatogram and the CE-LIF electropherogram of the IgG *N*-glycans was

achieved by the combination with offline weak anion exchange fractionation and targeted exoglycosidase digests. In the first dimension, the IgG oligosaccharide pool was separated based upon the degree of sialylation present, resulting in neutral, mono- (S1) and disialylated (S2) fractions. Each fraction was then profiled by UPLC-fluorescence and CE-LIF. Because of the high negative charge of APTS, fractionation of the IgG *N*-glycans for subsequent CE-LIF profiling was performed in their underivatized form, based on the retention time windows of the separation of the 2-AB labeled glycan pool. Thus, the use of UPLC-fluorescence in the second dimension is more facilitated, whereas for CE-LIF considerably more sample handling is required. However, the hyphenation with offline weak anion exchange fractionation is a powerful technique for the structural elucidation of sialylated glycan containing peaks. For highly sialylated biopharmaceuticals or glycoproteins other than IgG, the use of the approach introduced herein is highly recommended.

Figure 22 (A) displays the chromatogram for the anion exchange separation of the 2-AB labeled IgG *N*-glycan pool operated in semi-preparative mode. The individual profiling of the collected fractions using UPLC-fluorescence is depicted in Figure 22 (B). Potential overlaps between differently charged structures are well resolved in the neutral, S1 and S2 traces. Moreover, the removal of sialic acid upon enzymatic treatment of the individual charged fractions resulted in peak shifts to unoccupied regions of the chromatogram or electropherogram, thus allowing direct alignment with the previously annotated neutral fraction. Following sialidase digestion, as also displayed in Figure 22 (B), the identity of the oligosaccharides present was established leading to the structural annotation of the S1 trace. Interestingly, the main glycans present besides the predominant FA2G2S1 and its counterpart containing a bisecting GlcNAc residue peaks were found to correspond to a long armed FA2[3]G1S1 glycan (19.2% of the total peak area of fraction S1) and its bisecting GlcNAc carrying analogue. However, only minor quantities of the other monogalactosylated S1 isomer were present (1.2% of the total peak area of fraction S1). The higher quantities of the monosialylated FA2[3]G1 isomer were particularly surprising as for neutral IgG glycans significantly lower quantities of this isomer were observed. The α1-6 linked mannose residue of the trimannosyl core is reported to be involved in multiple interactions with the polypeptide structure in the $C_H2$ domain, whilst the antenna extending from the α1-3 mannose residue extends into the interstitial space between the heavy chains[19]. Therefore, sialylation of the α1-3 antennary GlcNAc residue is more favorable due to a combination of steric availability and also minimal interaction with the polypeptide chain that appears to play a governing role in the processing of the α1-6 antenna[22, 26]. Furthermore, it has been demonstrated that due to the confined conformational space between the heavy chains, combinations of sialylation leading to no more than two sialic acids present in the Fc region are the most

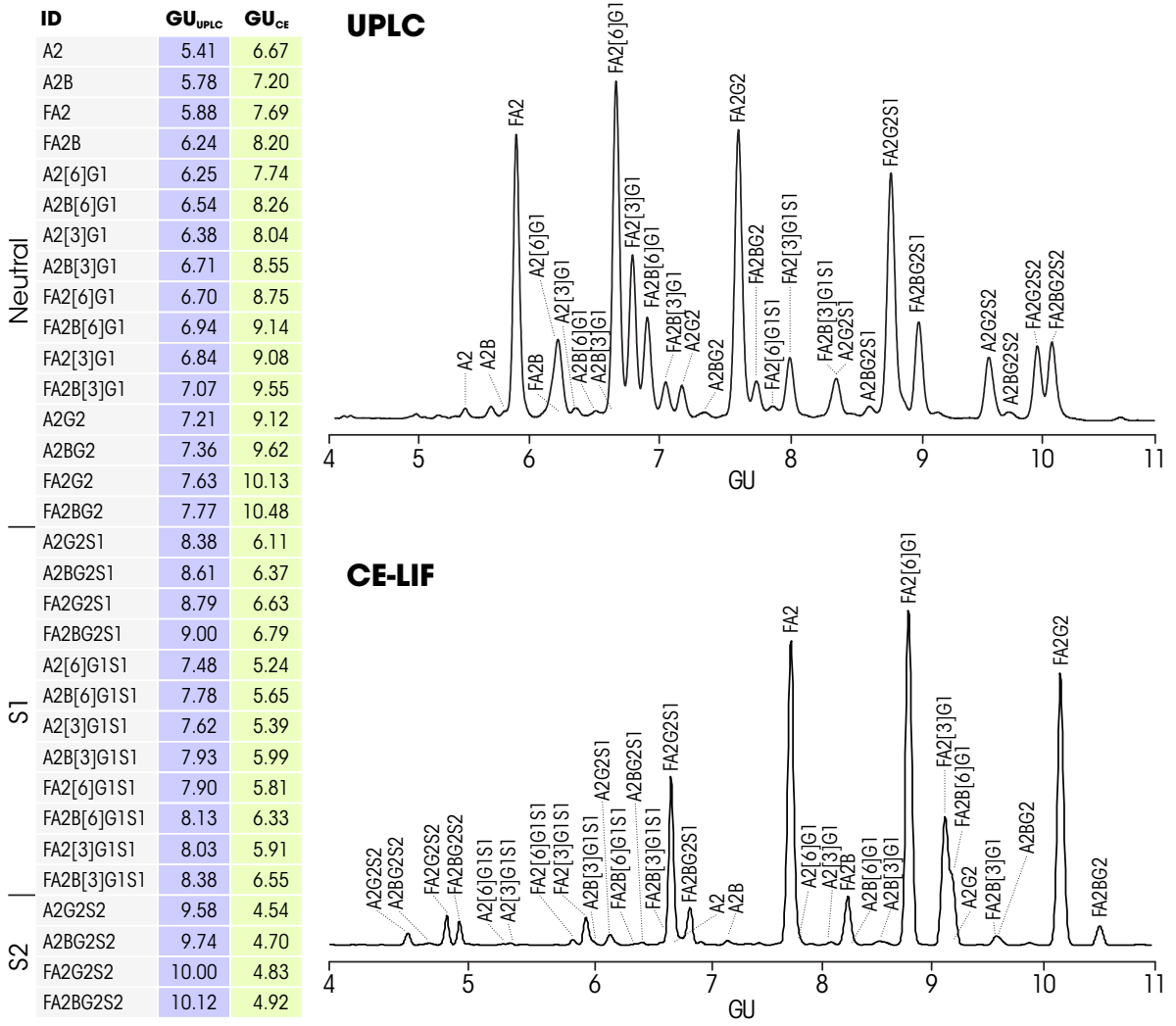
stable[236-237]. Heterogeneous pairing of neutral and S2 structures and homogenous pairing of monosialo monogalactosylated antennary type structures extending from the $\alpha$1-3 linked core mannose in the Fc region is therefore expected. Such combinations have been implicated with exhibiting the highest degree of anti-inflammatory activity[29, 236].

Figure 22 (C) and (D) show the aligned electropherograms of the S1 and S2 charged fractions with subsequent digestion by fucosidase, sialidase and a combination of the two. Fucosidase treatment of the intact sialoforms allowed direct visualization of the afucosylated monosialylated glycans. Comparing the peaks present after desialylation and subsequent defucosylation with the experimentally determined exoglycosidase digestion pathway in Figure 21 allowed for structural annotation.

### 4.4.4 Data combination: From comparison to a unified approach

CE-LIF displayed superior selectivity for the separation of sialylated structures, in particular, monosialylated glycans wherein it was possible to separate fucosylated and afucosylated structural analogues in a region of the electropherogram prior to the migration of all neutral structures. In the UPLC-fluorescence trace, the low abundant afucosylated monosialylated glycans (present on human IgG glycans were only visible following fucosidase digestion) and are lost under the more intense neutral peaks in the chromatogram corresponding to the total IgG *N*-glycan pool. For neutral glycans UPLC-fluorescence displayed superior selectivity over CE-LIF as can be visualized for the separation of the monogalactosylated positional isomer peaks and their analogues containing a bisecting *N*-acetyl glucosamine residue.

Exhaustive structural annotation of the peaks in the UPLC-fluorescence chromatogram and CE-LIF electropherogram of the IgG *N*-glycan pool was possible by combining the knowledge generated using our multidimensional approach. 32 *N*-glycan structures, expressed as experimentally determined mean GU values, are presented in the table inset in Figure 23. The 32 structures were distributed across 25 peaks separated by UPLC-fluorescence, and across 26 peaks separated by CE-LIF using a 20 minute run in each instance and the experimental conditions described previously. Comparison in a 2D plot, presented in Figure 24, revealed clustering of the individual oligosaccharides based upon their degree of sialylation due to the advantageous orthogonality introduced by the fundamental differences in the underlying separation mechanisms. This allowed for unsupervised charge based fractionation and grouping without the need for an additional separation dimension, thus highlighting the benefit of the combined use of both UPLC-fluorescence and CE-LIF.



| | ID | $GU_{UPLC}$ | $GU_{CE}$ |
|---|---|---|---|
| Neutral | A2 | 5.41 | 6.67 |
| | A2B | 5.78 | 7.20 |
| | FA2 | 5.88 | 7.69 |
| | FA2B | 6.24 | 8.20 |
| | A2[6]G1 | 6.25 | 7.74 |
| | A2B[6]G1 | 6.54 | 8.26 |
| | A2[3]G1 | 6.38 | 8.04 |
| | A2B[3]G1 | 6.71 | 8.55 |
| | FA2[6]G1 | 6.70 | 8.75 |
| | FA2B[6]G1 | 6.94 | 9.14 |
| | FA2[3]G1 | 6.84 | 9.08 |
| | FA2B[3]G1 | 7.07 | 9.55 |
| | A2G2 | 7.21 | 9.12 |
| | A2BG2 | 7.36 | 9.62 |
| | FA2G2 | 7.63 | 10.13 |
| | FA2BG2 | 7.77 | 10.48 |
| S1 | A2G2S1 | 8.38 | 6.11 |
| | A2BG2S1 | 8.61 | 6.37 |
| | FA2G2S1 | 8.79 | 6.63 |
| | FA2BG2S1 | 9.00 | 6.79 |
| | A2[6]G1S1 | 7.48 | 5.24 |
| | A2B[6]G1S1 | 7.78 | 5.65 |
| | A2[3]G1S1 | 7.62 | 5.39 |
| | A2B[3]G1S1 | 7.93 | 5.99 |
| | FA2[6]G1S1 | 7.90 | 5.81 |
| | FA2B[6]G1S1 | 8.13 | 6.33 |
| | FA2[3]G1S1 | 8.03 | 5.91 |
| | FA2B[3]G1S1 | 8.38 | 6.55 |
| S2 | A2G2S2 | 9.58 | 4.54 |
| | A2BG2S2 | 9.74 | 4.70 |
| | FA2G2S2 | 10.00 | 4.83 |
| | FA2BG2S2 | 10.12 | 4.92 |



***Figure 23:*** *Comprehensive annotation of UPLC-fluorescence and CE-LIF profiles of the N-glycans released from serum polyclonal IgG using the combination of the exoglycosidase digestion data and the weak anion exchange fractionation. Insets show a table of the experimentally determined GU value for each structure annotated in the UPLC-fluorescence and CE-LIF profiles. Adapted from*[233].

The precision of each technique was evaluated using triplicate injections for UPLC-fluorescence and n=6 for CE-LIF. GU value deviations were RSD < 0.1% for both UPLC-fluorescence and CE-LIF. Manual integration of peaks containing a single glycan structure across both the chromatogram and electropherogram returned matching quantitation for UPLC-fluorescence and CE-LIF. Deviations in the relative percentage area were on average < 2% for each of the integrated peaks.

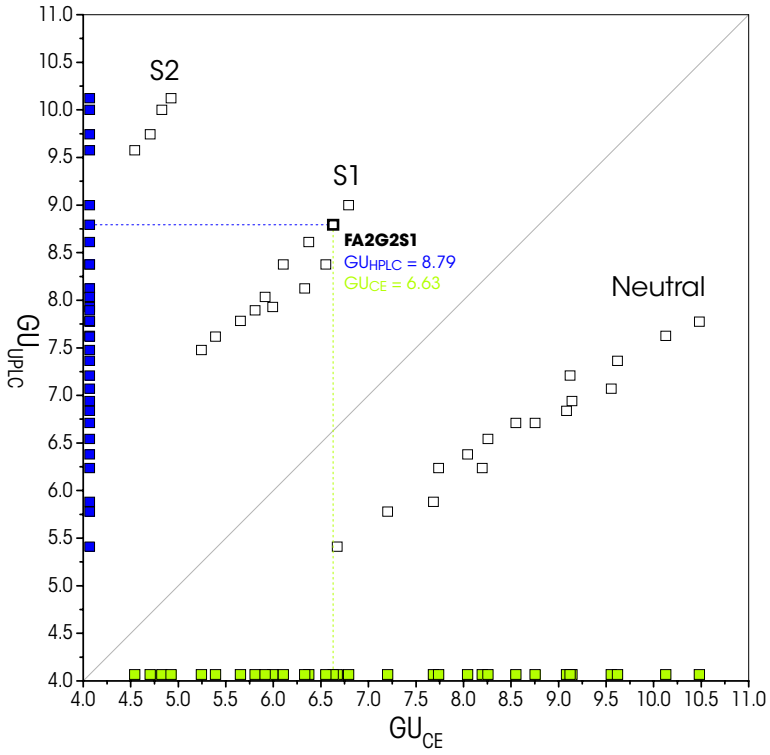



***Figure 24:*** *Comparative 2D plot of the experimentally determined GU value for each structure annotated in the UPLC-fluorescence and CE-LIF profiles. Adapted from*[233].

## 4.5 Identification of *N*-glycans bearing mannose-6-phosphate and their glyco-site

The acid hydrolase β-glucuronidase is a 74.7 kDa lysosomal enzyme, which exists as a homotetrameric structure, exhibiting four *N*-glycosylation sites on each monomer at asparagines 173, 272, 420 and 631. *N*-glycans released from β-glucuronidase *via* PNGase F were profiled using both UPLC-fluorescence on the 1.7 µm HILIC phase and CE-LIF with the previously established methods (respective sections 4.1 and 4.2).

### *4.5.1* *N*-Glycan profiling and exoglycosidase digestion identified peaks that may contain M6P residues

The β-glucuronidase *N*-glycan pool was digested with a panel of exoglycosidase enzymes (α2-3/6/8 sialidase, α1-2/3/4/6 fucosidase, β1-3/4 galactosidase and β-hexosaminidase) to structurally characterize *de novo* the glycosylation present. Particular emphasis was placed on the identification of peaks containing manose-6-phosphate residues. In Figure 25 the annotated *N*-glycan profiles are presented as experimentally determined using both UPLC-fluorescence and CE-LIF based upon the exoglycosidase digestion pathways and also the use of anion exchange chromatography based fractionation. The most abundant

oligosaccharides present on β-glucuronidase were a core fucosylated (F) mono antennary (A1) monogalactosylated (G1) glycan FA1G1, oligomannose 5 ($Man_5GlcNAc_2$), oligomannose 9 ($Man_9GlcNAc_2$) and a core fucosylated (F) bi-antennary (A2) di-galactosylated (G2) glycan FA2G2 corresponding to 18%, 16%, 15% and 14% of the total *N*-glycan pool, respectively. For the other complex-type oligosaccharides present, mono- and bi-antennary glycans with or without an α1-6 linked fucose on the reducing terminal *N*-acetyl glucosamine (GlcNAc) residue, variable antennary galactosylation which in some instances was further processed to display one terminal sialic acid represented the major portion of such glycans. Tri- and tetra-antennary glycans were present at much lower levels and again when sialylated, a maximum of only one terminal sialic acid was detected. Exoglycosidase digestions indicated a number of peaks in Figure 25 (marked with asterisks) that did not change either their elution or migration positions following enzymatic treatment, potentially due to the presence of non-reducing terminal M6P residues.

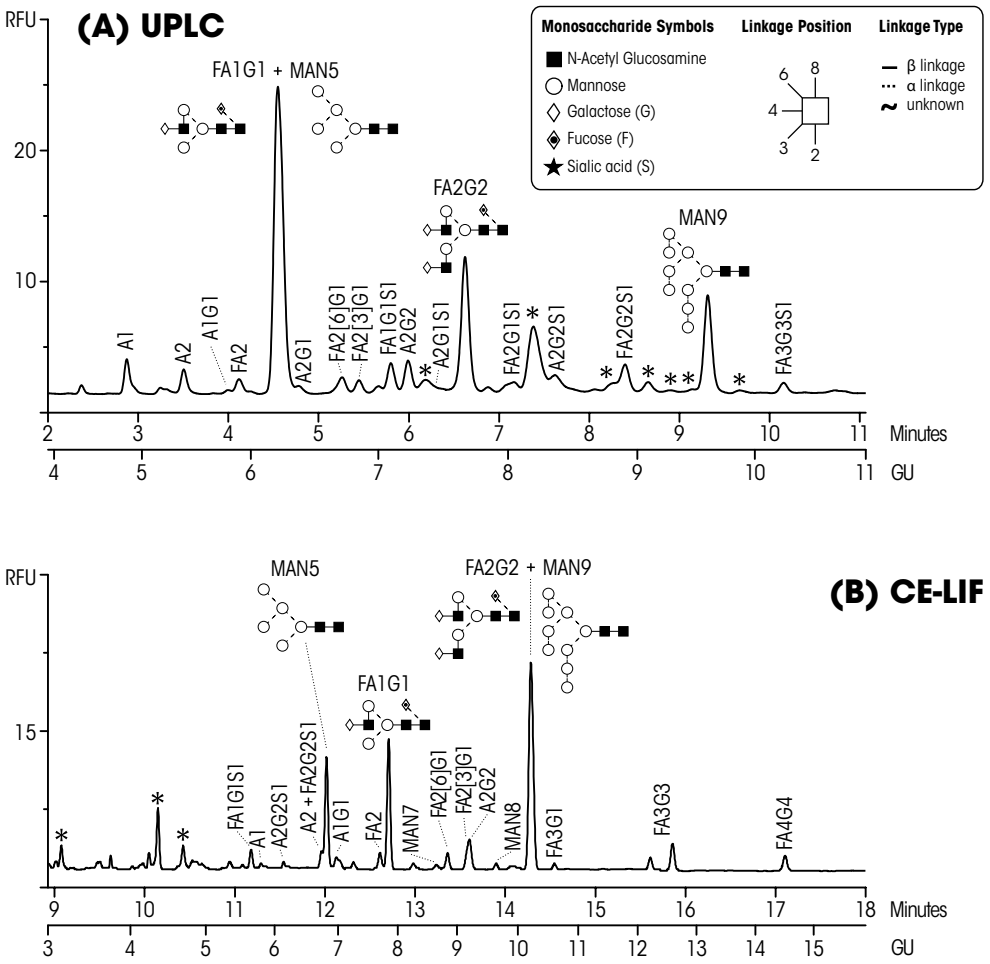



***Figure 25:*** *Analysis of the N-glycosylation present on recombinant β-glucuronidase using both UPLC-fluorescence and CE-LIF. Structural annotation was performed using a combination of the data generated by digestion using an exoglycosidase enzyme array and also weak anion exchange fractionation. Peaks marked with an asterisk were resistant to exoglycosidase digestion and are therefore suspected to contain M6P residues. Adapted from[238].*

### 4.5.2 AEC retention indicates the presence of negative charges from a source other than sialic acid

The *N*-glycan pool from β-glucuronidase was also separated by AEC on a weak anion exchanger. The resulting chromatogram was compared with that recorded following the injection of 2-AB labeled *N*-glycans from bovine fetuin, commonly used as a standard for the annotation of the degree of sialylation present. In Figure 26 (A), two retained peaks were present in the β-glucuronidase AEC chromatogram, labeled charged fraction one and two. Charged fraction one corresponded to the

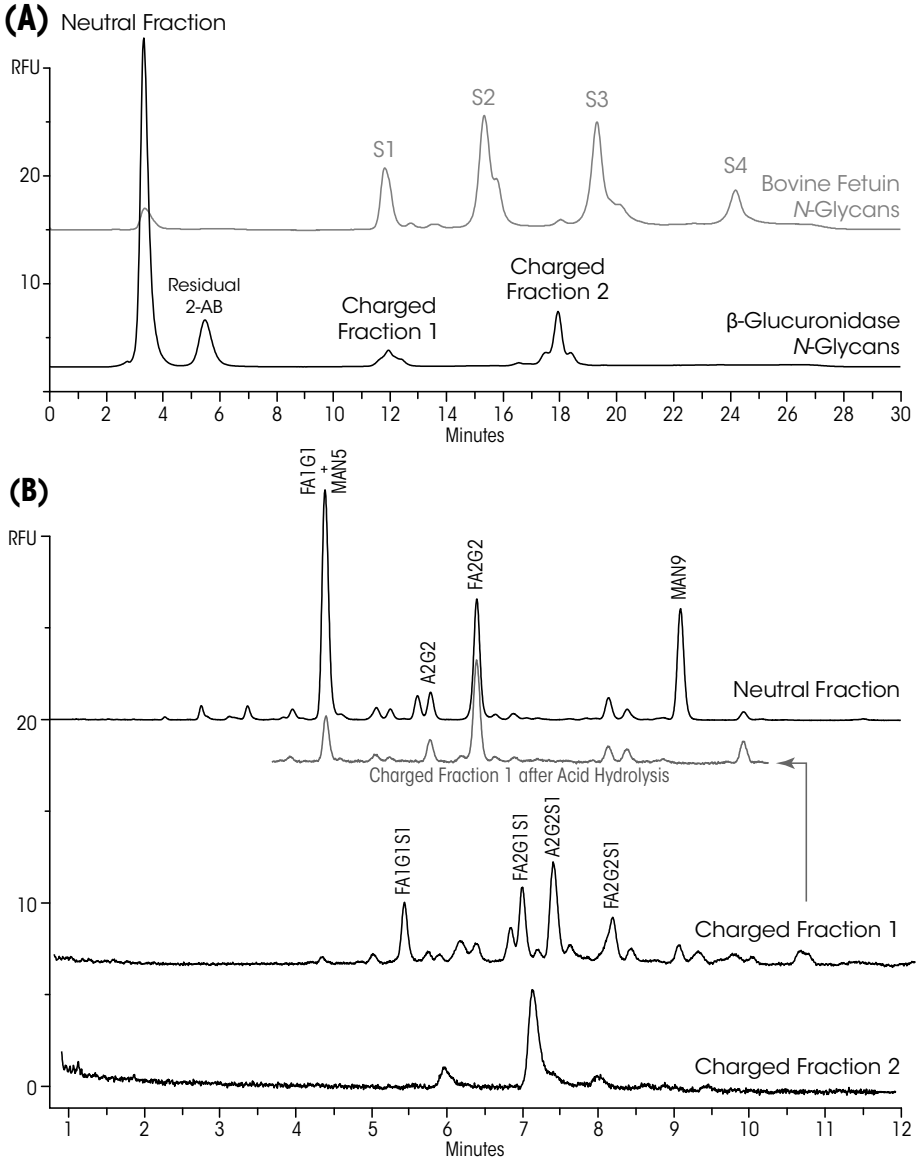



***Figure 26:*** *(A) Separation of the enzymatically liberated N-glycan pool of recombinant β-glucuronidase using weak anion exchange fractionation overlaid with the separation of N-glycans from bovine fetuin used as a standard for assessment of the degree of sialylation. (B) UPLC-fluorescence chromatograms for each of the collected fractions, inset is the trace following weak acid hydrolysis of the first charged fraction confirming that retention in this instance occurs due to the presence of a single terminal sialic acid residue on each of these oligosaccharides. Adapted from*[238].

position of the monosialo glycans present in the fetuin *N*-glycan pool and agreed with the previous exoglycosidase digestion data. Charged fraction two eluted at ~18 minutes in a region of the chromatogram that did not correspond to the peaks in the fetuin AEC *N*-glycan trace. Each fraction was collected and profiled by UPLC-fluorescence as shown in Figure 26 (B). Weak acid hydrolysis for sialic acid removal was also performed. As shown in Figure 26 (B), the retention of the first charged fraction on the weak anion exchanger was found to be due to the presence of one sialic acid as indicated on each of the *N*-glycans present. The glycans present in the second charged fraction did not contain sialic acid suggesting the presence of a charged substituent other than sialic acid. Moreover, when separated by UPLC-fluorescence the resulting peaks were found to correspond to those that were resistant to exoglycosidase digestion as indicated in Figure 25 again suggesting the presence of M6P on the *N*-glycans present.

### 4.5.3 Selective MOAP enrichment and alkaline phosphatase digestion

To further investigate the presence of phosphorylated *N*-glycans as indicated by the exoglycosidase digestion and AEC fractionation, mixed oxide affinity purification (MOAP) using a mixed $TiO_2/ZrO_2$ sorbent was performed to selectively enrich M6P containing oligosaccharides in the β-glucuronidase *N*-glycan pool. A mixed $TiO_2/ZrO_2$ sorbent was chosen to minimize any potential experimental bias and maximize the enrichment of oligosaccharides displaying either mono- or *bis*-M6P. Both the flow through and eluate were individually collected and analyzed by UPLC-fluorescence as shown in Figure 27 (A, traces ii & iii) compared with the total β-glucuronidase *N*-glycan pool in Figure 27 (A trace i). Upon examination of the resulting chromatogram for the MOAP enriched fraction it was noted that the trace contained the same peaks as were present in the chromatogram obtained following UPLC-fluorescence profiling of the second AEC charged fraction, Figure 26 (B). This selective enrichment of these particular oligosaccharides using MOAP was consistent with the presence of M6P residues as retention by MOAP occurs due to the interaction between phosphate present and the metal oxide media. To further test this, the β-glucuronidase *N*-glycan pool was treated with alkaline phosphatase and following overnight digestion profiled by UPLC-fluorescence and CE-LIF. The resulting trace is depicted in Figure 27 (B).

Following alkaline phosphatase digestion two peaks were observed to appear in the resulting UPLC-fluorescence and CE-LIF traces which were identified to correspond to $Man_6GlcNAc_2$ and $Man_7GlcNAc_2$. The difference in UPLC-fluorescence glucose unit (GU) value between $Man_6GlcNAc_2$ and $Man_6GlcNAc_2$-M6P corresponding to the presence of phosphate was determined to be ~1.3 GU units due to the increased hydrophillicity of the phosphorylated oligosaccharide.

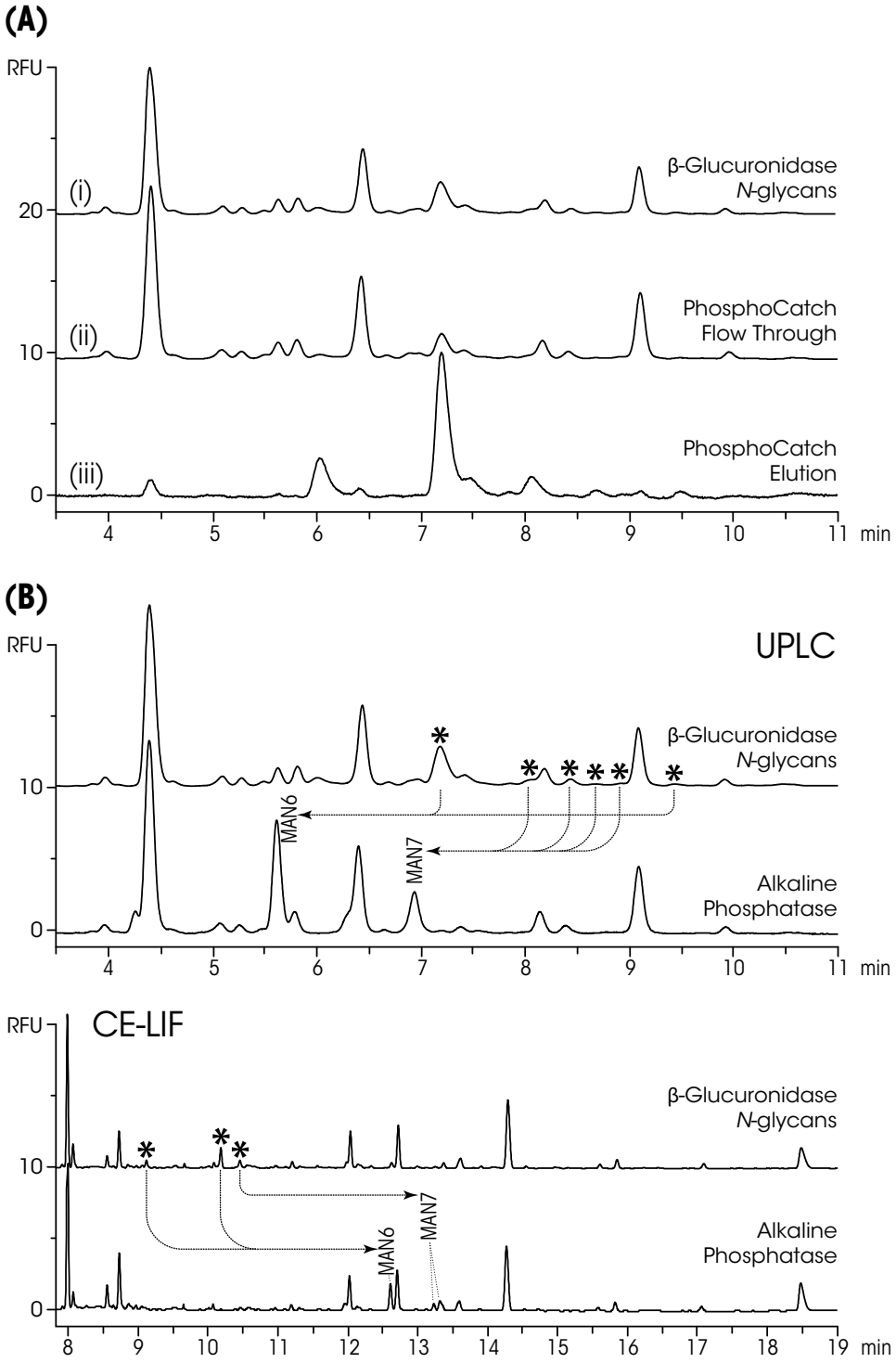



***Figure 27:*** *(A) Normalized UPLC-fluorescence chromatograms for β-glucuronidase N-glycans both before (trace i) and following fractionation using the $TiO_2/ZrO_2$ mixed oxide affinity media (trace ii flow through, trace iii eluate). (B) Resulting UPLC-fluorescence and CE-LIF traces following alkaline phosphatase digestion confirming the presence of phosphorylated oligosaccharides and identifying them as $Man_6GlcNAc_2$ and $Man_7GlcNAc_2$. Adapted from*[238].

Inspection of the resulting fluorescence chromatograms for the undigested β-glucuronidase *N*-glycan pool and the MOAP enriched fraction showed a number of low intensity broad peaks eluting between eight and eleven minutes that moved upon alkaline phosphatase digestion. These peaks are consistent with the presence of the different positional isomers for $Man_7GlcNAc_2$ that occur depending upon the location of attachment of the α1-2 linked mannose residues.

A peak at GU 9.8 in the fluorescence chromatogram shown in Figure 25 and at GU 3.1 in the LIF electropherogram of the β-glucuronidase *N*-glycan pool, in both cases at ~1% of the total, was also observed to shift following alkaline phosphatase digestion. Based upon the retention and migration behavior of this peak the presence of *bis*-M6P is suspected on $Man_6GlcNAc_2$.

### 4.5.4 Site-specific analysis reveals presence of M6P at asparagine 272, 420 and 631.

Having identified the presence of M6P containing glycans in the total *N*-linked oligosaccharide pool of recombinant β-glucuronidase, a site-specific glycosylation analysis was then performed to determine which of the four *N*-glycosylation sites display the M6P residue on their attached glycans. Two separate analytical approaches were undertaken, a peptide mapping centered approach to identify glycosylation sites carrying M6P tagged glycans and then a MOAP enrichment of phosphorylated glycopeptides with subsequent LC-MS/MS analysis to assess the degree of phosphoglycan occupancy.

*In silico* digestion of β-glucuronidase with a variety of proteases was performed using the Swiss Institute of Bioinformatics Peptide Cutter and Peptide Mass tools in order to identify the protease or combination of proteases capable of generating peptides of appropriate length carrying a single *N*-glycosylation site (Swiss Institute of Bioinformatics, Peptide Cutter tool: http://www.expasy.ch/tools/peptidecutter/ and Peptide Mass tool: http://www.expasy.ch/cgi-bin/peptidemass.pl. The accession number for β-glucuronidase used in all informatics analyses was P08236). From these evaluations both Asp-N and Lys-C were individually employed for β-glucuronidase digestion. The use of the two proteases was investigated to maximize the glycoprotein coverage in all subsequent glycomics investigations. To simplify the resulting peptide maps as shown in Figure 28 (A), glycopeptides present following digestion were enriched using microscale HILIC fractionation and separated by reverse-phase liquid chromatography. Each peak was individually collected, treated with PNGase F to remove *N*-glycans if present and then analyzed by LC-MS/MS to identify the peptide.

Using the peptide mapping approach with subsequent UPLC-fluorescence *N*-glycan profiling, trace levels of $Man_7GlcNAc_2$-M6P glycans at asparagine 272 and 420 in both the Asp-N and Lys-C glycopeptide

fractions and $Man_6GlcNAc_2$-M6P glycans at asparagine 631 present in the Asp-N digest only were identified, Figure 28 (B, C and D).

To assess the degree of phosphoglycan occupancy at asparagine 272, 420 and 631, MOAP enrichment of β-glucuronidase Asp-N glycopeptides was performed. Both the flow through fraction and the eluate were individually collected and analyzed by LC-MS/MS following Glu-C digestion, 25 mM phosphate buffer pH 7.8, 25°C overnight. Peptide fragments from the Asp-N peptide containing asparagine 631 were detected in both the flow through and the MOAP eluate. Peptide fragments from the Asp-N peptides containing asparagine 272 or 420 were not detected, most probably due to their small size and high proportions of basic residues, thereby creating the possibility for high charge states and low m/z ratios that may not be selected for MS/MS during data dependent analysis. Extracted ion chromatograms (EICs) were generated in order to assess the ratio of phosphorylated to non-phosphorylated glycopeptide present, (m/z 1546.77947, retention time 21.3 minutes, peptide sequence $^{633}$TRYPHSVAKSQCLE$^{646}$). Based upon the peak area in the EICs it is estimated that approximately 8.1 % of all asparagine 631 present displays glycans bearing M6P.

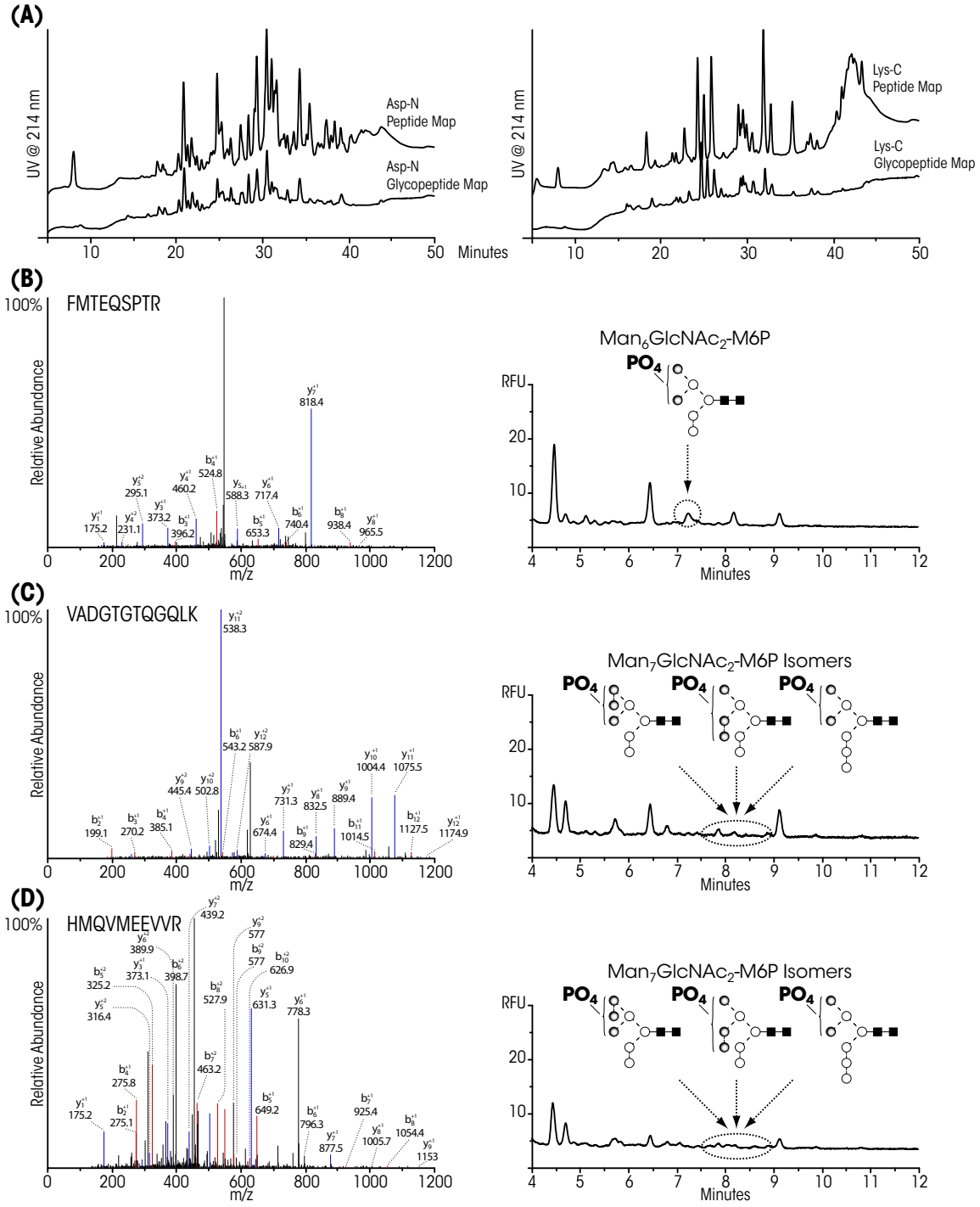



***Figure 28:*** *(A) Reverse-phase separation of β-glucuronidase peptides and HILIC enriched glycopeptides following digestion with either Asp-N or Lys-C. Identification of the peptide fragment corresponding to: (B) the Asp-N peptide carrying asparagine 631 along with the presence of $Man_6GlcNAc_2$-M6P in the resulting HILIC-UPLC-fluorescence trace, (C) the Asp-N peptide carrying asparagine 272 along with the presence of $Man_7GlcNAc_2$-M6P in the resulting HILIC-UPLC-fluorescence trace and (D) the peptide fragment corresponding to the Asp-N peptide carrying asparagine 420 along with the presence of $Man_7GlcNAc_2$-M6P in the resulting HILIC-UPLC-fluorescence trace. In all instances a shaded mannose symbol represents a potential site of phosphate attachment. Adapted from*[238].

# CHAPTER 5

# Discussion

## 5.1 Identification of cancer associated serum *N*-glycome changes by UPLC

The advent of commercially available sub-2 μm HILIC stationary phases represents a major advance for practitioners of liquid chromatography based glycomics. The attractive attributes of HILIC based glycan separations include predictive selectivity where retention increases with glycan size. The major limitation to date has been the relatively low efficiency of the available HILIC stationary phases. For example, Lauc and co-workers have published a number of HILIC-fluorescence based studies examining the genetic regulation of the plasma *N*-glycome with the maximum number of distinct chromatographic peaks reported for integration and subsequent analysis being sixteen using a one hour gradient method[239-242]. The sub-2 μm HILIC stationary phase operated by ultra performance liquid chromatography provides a considerable increase in chromatographic efficiency. Feature numbers, *i.e.* the number of distinct and integratable peaks were determined to be 30, 46 and 53 which represents an increase in feature numbers of 1.9, 2.9 and 3.3 fold as compared to those of Lauc and co-workers for the 5, 10 and 15 cm analytical columns, respectively. The improvement in selectivity was also recorded using significantly reduced analytical run times of 10, 20 and 30 minutes for the 5, 10 and 15 cm analytical columns which permits a 6, 3 and 2 fold increase in sample throughput, respectively as opposed to the previously mentioned 60 minutes required when using HPLC.

The increase in chromatographic efficiency permits the separation of conformational and linkage specific isomeric oligosaccharides and also glycans containing a bisecting *N*-acetyl glucosamine residue. The ability to resolve and reliably determine lower abundant and biologically significant oligosaccharide species represents a considerable advance in glycomics separations. This is particularly relevant when searching for cancer associated alteration in glycosylation or attempting to gain an understanding into the physiology of the disease. The activity of the glycosyltransferase involved in the attachment of the bisecting *N*-acetyl glucosamine residue on *N*-glycans, *N*-acetyl glucosaminyl transferase III (GnT-III), has been shown to suppress cancer metastasis due to stabilization of the cell surface adhesion molecule E-cadherin resulting in the promotion of homophilic adherin based cell adhesion[243-244]. In addition to this, the presence of *N*-glycans containing the bisecting *N*-acetyl

glucosamine residue on IgG has also been shown to increase antibody dependent cellular cytotoxicity activity (ADCC) because the bisecting monosaccharide inhibits core fucosylation and increases affinity of IgG molecules for the Fcγ receptor III[244].

Statistical analysis of the alterations in the total serum *N*-glycome in the course of disease progression revealed lower levels of asialo- and monosialylated core fucosylated glycans, normally carried on IgG molecules which are present in serum at concentrations of ~10 mg/mL[47]. Also an increase in branching and sialylation of *N*-glycans was observed, that represents a commonly reported hallmark of cancer[245-246]. However, the problems of using serum as a biological matrix for biomarker discovery have been well documented[44]. Despite the reported obstacles, serum remains the biofluid of choice due to its ease of availability. The individual contributions of *N*-glycans derived from IgG, haptoglobin, transferrin and α1-acid glycoprotein to total serum trace are particularly relevant for studies focusing upon crude serum glycomics. The present data suggests that changes recorded in the total serum glycome based upon the alterations in the levels of individual peaks are mostly likely due to alterations present in the glycosylation present on one of these four highly abundant glycoproteins. Although haptoglobin, transferrin and α1-acid glycoprotein are abundant serum glycoproteins and their individual glycans appear to populate the majority of the serum glycome, the presence of the same chromatographic peaks after immunoaffinity depletion indicates the presence of these biologically significant glycans also on other lower abundant serum proteins. Despite the reported limitations of serum it appears that for glycomics, due to the limited numbers of possible oligosaccharides that can be present as orchestrated by the glycosylation machinery of the system under study, serum is still a biological matrix with great potential for biomarker discovery.

Two-dimensional electrophoresis and affinity purification enabled effective partitioning of the pathologically staged cancer serum and individual analysis of the alterations in the glycosylation present on IgG, haptoglobin, transferrin, and α1-acid glycoprotein as possible markers for stomach cancer progression. Increased levels of agalactosyl *N*-glycans on IgG with disease progression can be associated with previously reported pro-inflammatory properties of IgG isoforms carrying agalactosyl *N*-glycans which, when clustered are a ligand for mannose binding lectin and complement activation[247]. Furthermore, in accordance with previously mentioned biomarker discovery studies covering many other cancers, the levels of sialylated glycans on haptoglobin, transferrin and α1-acid glycoprotein revealed an increase with disease progression[68].

In conclusion, the applicability of newly available sub-2 μm HILIC stationary phases for ultra performance liquid chromatography based glycomics has been demonstrated. Both a significant increase in

efficiency and decrease in analysis time with improvement in the selectivity of *N*-linked oligosaccharide separations compared to previous HPLC based studies were presented. UPLC-fluorescence profiling of serum from pathological patient pools permitted the identification of significant cancer associated alterations in the serum glycome of patients bearing stomach adenocarcinoma. The contribution of highly abundant serum glycoproteins to the total serum *N*-glycome was also determined and their individual alterations in the glycosylation with increasing disease pathogenesis were evaluated. The novel application of UPLC-fluorescence using the sub-2 μm BEH Glycan HILIC phase as presented herein may act as the basis for further serum glycomic biomarker discovery experiments in the future.

## 5.2 CE reflects molecular conformation and configuration of oligosaccharides

The development of a practical method for carbohydrate profiling by capillary electrophoresis including eligible normalization approaches is crucial to ensure experimental precision and compare the delicate migration differences of structurally similar or isomeric species. The introduction of a normalized scale such as glucose units instead of migration time increases experimental precision when both ladder and sample traces are aligned based upon appropriate bracketing standards. Although, the comparison between the results of the two GU calculation methods revealed only minor differences, a step-by-step linear interpolation as used in eq. 4 (Theoretical Background section) was considered a less accurate approximation compared to piecewise- (spline interpolation) or continuous higher order polynomial functions. Another advantage of a continuous polynomial approach is its easy potential for automation wherein solving a single polynomial function directly provides the sugar unit (SU) values. This is of particular interest when SU values need to be calculated for multiple peaks, potentially even in numerous (repetitive) runs, *i.e.*, the software does not have to locate preceding and following ladder standard peaks for each peak of interest.

The described regression analysis discrepancies between malto- and cellooligosaccharides (DP 1-6) were due to significant differences in the molecular shape of the two homooligosaccharide ladders, originated from their distinct anomeric configuration. The more flexible equatorially oriented bond in the β-anomeric configuration leads to extended rod-like structures, whereas the axial α1→4 linkage is bent intending to form helical structures supported by intramolecular H-bonds[248-249]. Energy minimized $Glc(\alpha1\rightarrow4)_{12}$ and $Glc(\beta1\rightarrow4)_{8}$ oligosaccharide models, depicted in Figure 12, also exhibited a helical and an extended rod-like shape, respectively. Faster migration, thus concomitantly anticipated smaller hydrodynamic radius, of DP 2, 3, 4 and 5 cello- compared to the corresponding maltooligosaccharides was not suggested in previously reported results based on size-exclusion chromatography (SEC)[249]. The faster migration of small cellooligosaccharides could be caused by the decreased

electromigration resistance of the linear-like β1→4 configuration in contrast to the bent α1→4 linked oligosaccharides. With increasing size, intra-molecular hydrogen bonds between the hydroxyl groups of the glucose residues and bridging water molecules stabilize the helical conformation of maltooligosaccharides, thus rendering α1→4 linked oligo-glucoses to form a more compact and less flexible helical structure than that of its linear β1→4 linked counterpart[250]. The increasing distances between consecutively migrating β1→4 linked oligo-glucoses as well as the fact that maltohexaose already migrated faster than cellohexaose were, however, in agreement with earlier size exclusion chromatography results[249]. Also the curve shaped increase for both malto- and isomaltosides in DP *versus* migration time, reflecting the relative changes in hydrodynamic radius with increasing DP, was in accordance with previously reported results based on SEC[251]. For maltooligosaccharides, this curve flattened after 7-8 glucose units, which was originated from their tendency to form a full helical turn upon reaching DP 6[252-253]. Due to the fluorescent labeling reaction *via* reductive amination, the glucose at the reducing termini was present in its open form and consequently influenced the unit numbers required to close the first helical turn. Hence, the number of necessary monosaccharides to form a full helix was expected to be ≥7. The extra dihedral angle within the α1→6 glycosidic linkage and the reportedly associated increased flexibility[254] did not cause any decisive differences in the electromigration behavior in the smaller range (DP 1-6) of isomaltosides compared to maltosides.

All four homooligosaccharide ladders revealed a slope bend in log(μ) *versus* log(M) plots, thus supported the assumption of changes in their geometrical configuration with increasing degree of polymerization. Slope values in the range of -0.62 to -0.68 for α1→4 linked maltosides further endorsed the assumption of the formation of a relaxed helical structure in solution[250]. The different slope value of -0.62 in the lower molecular size range (DP 1-9) was expected due to incompletion of the first full turn of the helix but also due to the impact of the open ring at the reducing terminal monosaccharide as a result of the labeling reaction. The lower slope values in the higher DP range of isomaltosides (DP 9-24) and β1→4-linked oligo-glucoses (DP 3-6) suggested more extended and rather linear/bulky structures, respectively.

Potentially associated changes in the molecular conformation of α1→4 and α1→6 linked oligosaccharides with increasing temperature revealed linear changes in mobility and only minor differences in the presented slope values. Therefore, it can be concluded that the additional thermal energy did not have an effect on the geometrical properties of malto- and isomaltooligosaccharides or on their stability in the investigated temperature range.

In HILIC experiments, isomaltooligosaccharides revealed more pronounced presentation of hydroxyl groups that are not shielded by hydrogen bridges, because the α1→6 linkage is characterized by absent

hydrogen bonds between adjacent residues, thus rendering this configuration more flexible and amenable to hydration[251, 255]. In contrast, the intra-molecular bridges in α1→4 linked maltooligosaccharides, lead to the formation of a helical conformation, decrease the availability of hydroxyl groups and subsequently lead to decreased retention. Due to the orthogonal separation mechanism in HILIC, geometrical changes towards the completion of a full helical turn in maltohexaose or -heptaose were not as apparent as during the CE studies.

In conclusion, carbohydrate positional and/or linkage isomers, bearing the same charge-to-mass ratio even at equal molecular masses, can be readily separated by CE due to its inherent separation principle based on hydrodynamic volume equivalent radius differences. The application of sugar unit values (SU), rather than migration time or electrophoretic mobility, corrects for discrepancies possibly caused by separation temperature changes, column history, buffer composition discrepancies, etc., thus rendering it to a unified and comparable value. This enables the analysis of known carbohydrate structures (*e.g.* commercially available standards), determination of their sugar unit values and consecutive direct comparison to those of unknown sample peaks for easier structural elucidation. Besides tentative structural determinations, this approach eventually leads to the accrual of a large number of SU values; and subsequently the establishment of a SU database enabling direct deduction of carbohydrate structural information from experimental CE data of complex unknown carbohydrate mixtures. At present, only the HILIC-HPLC approach is associated with available databases for searching structures assigned relative to retention times, represented as glucose units (GU)[86]. Certainly, an additional CE database will be a powerful tool to complement structural elucidation endeavors of unknown glycans. As for all such databases, standardization criteria need to be closely followed; *i.e.* the very same experimental conditions need to be applied under which the database was built, or otherwise structural assignment based on SU values will be inaccurate. A crucial part of the necessary experimental conditions is the usage of the adequate type of homooligosaccharide ladder standard. While in capillary electrophoresis, the application of a maltooligosaccharide ladder (α1→4 linked oligo-glucoses) as standard is prevalent, α1→6 linked isomaltooligsaccharides (also referred to as dextran) are the most commonly used standards for GU calculation in carbohydrate analysis by HPLC. As experimentally illustrated, the mere difference in glycosidic linkage type or anomericity can have a strong impact on the separation behavior of oligosaccharides and therefore, the choice of the appropriate ladder standard is crucial for the application of carbohydrate databases and the associated structure elucidation potential.

## 5.3 Aptitude of CE to reveal pulmonary disease associated glycan features

The positive acute phase protein haptoglobin possesses four potential *N*-glycosylation sites in its β-chain and was previously investigated by various liquid chromatographic[256] and mass spectrometric techniques[257]. It has been reported that haptoglobin *N*-glycans consist of predominantly bi- but also some tri- and tetra-antennary structures, which can be either core or antennary fucosylated[258]. The obtained structural identification by CE profiling in conjunction with targeted exoglycosidase digestions are in agreement with these findings and similar antennary distribution and fucosylation features were identified on a haptoglobin protein standard. High resolution of the CE method also proved valuable in the quantitative distinction between core and arm fucosylation and associated potential consideration in molecular diagnostics. However, due to the small scale patient cohort in the present study, attempts of biomedical conclusions are highly limited and only assumptions upon disease associated impacts on differential haptoglobin glycosylation are possible.

In conclusion, the presented experimental results revealed differential glycosylation in the inflammatory and malignant lung diseases that encourages continuing the quest for finding disease specific glycosylation markers. The current study, however, rather represents a proof of concept approach wherein the potential of CE-LIF for the robust profiling of complex glycans pools including positional and/or linkage isomers (*e.g.* core or antennary fucosylation) of biomedical interest is demonstrated. Gender and age variability of the *N*-glycans present on serum glycoproteins[259] require large scale studies in order to claim biomarker or diagnostic conclusions.

## 5.4 Unified use of CE and UPLC for rapid and confident IgG glycan analysis

Determining the structure/function relationship of the *N*-glycosylation on IgG requires comprehensive characterization by high performance analytical methods. Liquid phase chromatographic and electrophoretic separation techniques have found widespread application for the characterization of IgG *N*-glycans due to their ability to separate critical features such as positional isomers and peak area based quantitation[106, 135, 260-262].

As demonstrated here, using the applied buffer system, CE-LIF is outperformed by the 1.7 μm HILIC phase for the separation of neutral glycans and their afucosyl and bisecting GlcNAc containing analogues. Modification of the CE-LIF buffer system with additives that employ secondary equilibria enhancement of the selectivity would result in improved separations of such glycans[116, 263]. It is also acknowledged that bisecting GlcNAc containing glycans will not be present on monoclonal antibodies

produces in CHO cells due to the absence of the GnTIII gene[264-265]. However, such production systems are likely to express high mannose type *N*-glycans due to incomplete glycan processing due to the stress placed on the glycosylation machinery during production[206, 266]. Although not presented in the current study, the separation of high mannose glycans from bovine ribonuclease B overlaid with IgG *N*-glycans was investigated. No co-elution of the major mannose-5, 6 or 9 peaks was observed with any of the IgG *N*-glycan peaks for UPLC-fluorescence. In CE-LIF, the identification of the mannose-9 species necessitated an additional exoglycosidase digest step due to co-migration. The presented methodology can therefore be directly applied to high-throughput yet comprehensive glycosylation analysis of monoclonal antibodies for therapeutic or diagnostic purposes.

Out of a biological maximum of 36 possible structures based upon all possible combinations of fucosylation, galactosylation, sialylation and the presence or absence of the bisecting GlcNAc residue, finding of no more than 32 structures were reported previously[267]. The remaining 4 glycans are bigalactosylated monosialylated positional isomers, carrying a fucose and/or a bisecting GlcNAc residue. Using both UPLC-fluorescence and CE-LIF it is currently not possible to separate these monosialylated positional isomers on the bigalactosyl glycans due to structural similarity. For the monogalactosyl species it was possible to separate these positional isomers, however in those instances there was a difference of two sugar units, sialyl-galactose on one or the other antenna. This two unit difference generated sufficient geometric difference to enable the separation. Furthermore, preferential sialylation of the galactose residue present on the α1-3 antenna has been reported previously[268]. We found low abundance of the sialyl-galactose on the antenna extending from the α1-6 antennary mannose as opposed to the α1-3 antennary mannose. Therefore, the lack of ability to separate the monosialo bigalactosyl positional isomers may also be confounded by possible dynamic range issues leading to an adverse effect on their overall separation in both UPLC and CE-LIF. Low quantities of neutral and sialylated afucosylated *N*-glycans in the IgG oligosaccharide pool hindered their identification and required enzymatic defucosylation of higher abundant species for confirmation of either their migration or elution position. The ability to identify and quantify such oligosaccharides is however, of biological relevance, particularly in the case of therapeutic antibodies as these afucosyl glycoforms have been demonstrated to increase ADCC function[21].

Annotation of all 32 potential structures in the IgG *N*-glycan pool using a combination of anion exchange fractionation and exoglycosidase digestion steps with subsequent profiling, as presented herein and reported previously, necessitates over 48 hours[267]. Upon combination of the data sets generated individually by UPLC and CE-LIF using a 2D approach, all 32 oligosaccharides present are well separated within the confined two-dimensional space. The use of glucose unit values following time based

normalization and the implementation of CE-LIF bracketing standards further increased injection-to-injection reproducibility by eliminating any associated experimental variance[156]. Combining this high precision with the beneficial orthogonality of the data sets allowed for exhaustive and confident structural annotation of the IgG *N*-glycan pool within only 20 minutes using the 2D space. Similar approaches have been previously evaluated for both two- and three-dimensional distributions of the chromatographic data with subsequent annotation by positional matching[269-270].

In conclusion, the combined use of both of the applied techniques is recommended for rapid and comprehensive characterization of glycosylation present on antibodies for therapeutic administration or in the biomedical setting in order to maximize the analytical return and permit translation of the resulting data into a meaningful biological interpretation. The unified analytical platform opens the possibility for vector based annotation of unknowns based upon their positioning within the multidimensional data cloud and furthermore allows for the statistical analysis of the vector based positional match. This expands the potential for a 'score' based confidence system in the resulting annotation as routinely used in other 'omics' data analysis.

## 5.5 Importance of mannose-6-phospate carrying *N*-glycans on β-glucuronidase

β-glucuronidase is a 74.7 kDa lysosomal acid hydrolase responsible for the catabolism of heparan, dermatan and chondroitin sulphates. It exists as a homotetramer with four *N*-glycosylation sites present on each monomer at asparagine 173, 272, 420 and 631. Deficiency in β-glucuronidase results in mucopolysaccharidosis type VII, also known as Sly syndrome, which causes central nervous system degeneration, mental retardation and fatality in severe cases[271]. In addition, mucopolysaccharidosis type VII is often accompanied by reduced immune function[272]. Administration of recombinant β-glucuronidase has been effectively used *in vitro* to treat mucopolysaccharidosis type VII. The incorporation of increased levels of M6P on β-glucuronidase has been demonstrated to be a more efficacious treatment due to higher uptake of the therapeutic protein by a wide range of tissues, thereby resulting in a more complete clinical response[273].

Considering the biological significance of M6P, it is surprising that a limited number of analytical methods for its determination exist in the literature. Acid hydrolysis with subsequent monosaccharide profiling offers the ability to quantify the levels of M6P present but provides no information regarding the structural identity of the oligosaccharide, or the position on the polypeptide, to which the M6P residue was attached[274-275]. Nuclear magnetic resonance spectroscopy and mass spectrometry have also been used to elucidate the structure of oligosaccharides and the position of phosphate attachment within the

glycan[276-279]. However, due to their isobaric masses, unambiguous differentiation between phosphorylation or sulphation present on the glycans has been problematic when using mass spectrometric analysis[280]. Liquid chromatographic profiling of 2-aminobenzoic acid (2-AA) labeled oligosaccharides from lysosomal acid lipase (LAL), expressed in *Pichia Pastoris* or Chinese hamster ovary (CHO) cells, was reported using long gradients on an amino stationary phase with online fluorescence and mass spectrometric detection[281]. Structural annotation of the peaks in the fluorescence chromatograms was performed using the corresponding *m/z* value. Increased M6P incorporation was observed on the CHO derived protein and the phosphorylation was displayed on $Man_7GlcNAc_2$ and $Man_6GlcNAc_2$ glycans[281]. LC-MS analysis of glycopeptides, from two commercial recombinant α-galactosidase A treatments for Fabry's disease, was applied to determine the presence of *N*-glycans displaying M6P. $Man_7GlcNAc_2$ and $Man_6GlcNAc_2$ were again reported as the oligosaccharides carrying the phosphorylation[282]. Despite their successful application, these methods require extensive analysis times making them unsuitable for use during upstream process optimization and monitoring.

The presence of M6P is known to block further processing of such tagged oligosaccharides in the Golgi by the glycosyltransferase enzymes[36]. In an analogous manner, the stationary behavior of those peaks was attributed to blocking the activity of the exoglycosidase enzymes by the potential presence of the M6P residue. Therefore, the use of exoglycosidase digestion permits the identification of possible M6P containing targets but it does not allow for their structural characterization. For the selective enrichment and identification of *N*-glycans carrying M6P, the use of mixed oxide affinity purification (MOAP) was evaluated. MOAP has previously been applied for the enrichment of phosphopeptides[283-285]. Selective enrichment occurs due to the pH dependent bidentate interaction between the phosphate group and the metal oxide material, *e.g.* $TiO_2$ or $ZrO_2$[286]. A mix of $TiO_2$ and $ZrO_2$ was employed here to minimize experimental bias and maximize the enrichment of phosphorylated glycans. It has previously been reported that $TiO_2$ and $ZrO_2$ have different affinities for mono- or diphosphorylated peptides[286].

As opposed to the identification of $Man_6GlcNA_2$-M6P, reconciliation of the UPLC-fluorescence data for phosphorylated $Man_7GlcNA_2$-M6P and its phosphatase digest products was more complicated due to the potential presence of positional isomers that occur depending upon the location of attachment of the α1-2 linked mannose residues. A further degree of complexity exists as the location of the phosphate residue has also been shown to be variable among the terminal mannose residues[287-288]. Therefore, a total of eight $Man_7GlcNAc_2$-M6P monoesters are possible which leads to the multiplicity in the resulting chromatogram. The presence of the phosphate residue also results in peak broadening potentially due to secondary anion exchange interactions with the amido-functionalised HILIC phase. The combined use of CE-LIF and UPLC-fluorescence is beneficial as each of the possible $Man_7GlcNAc_2$-M6P positional

isomers has the same mass to charge distribution resulting in a single early migrating peak in the electropherogram. The ability to cross correlate the data individually generated using both techniques increases the confidence of the oligosaccharide structural assignments.

Previous studies concerning the glycosylation of β-glucuronidase using site directed mutagenesis to delete glycosylation sites have shown that in addition to protein sorting and shuttling to the lysosome, the *N*-glycans play a crucial role in the assembly of the active tetrameric form of the enzyme[289]. From the experimental data only a small portion of glycans contained *bis*-M6P, the oligosaccharide that shows the highest affinity for both the cation-dependent and -independent M6P receptors[288]. Interestingly, significantly more glycans bearing a single phosphate group were observed. It has previously been reported that two appropriately spaced monophosphorylated glycans can act in unison as a high affinity MPR ligand indicating flexibility in MPR mediated acid hydrolase lysosomal trafficking[40]. The large quantities of oligomannose glycans detected were also observed during x-ray crystallographic studies of β-glucuronidase and were noted to be located at asparagine 173[290]. Although not normally seen in x-ray studies, the oligomannose glycans were thought to have reduced thermal mobility due to their role in crystal packing with neighboring tetramers[290].

The observed presence of mono-antennary glycans on β-glucuronidase represents another interesting observation and suggests either a low UDP-GlcNAc flux or a short residence time in the Golgi body during the CHO cell culture[291]. However, as the glycans are only required for the organization of the β-glucuronidase tetramer but not required for subsequent enzyme activity[289], these mono-antennary short chain oligosaccharides may be a cellular processing feature designed to minimize possible non-specific secondary interactions between the glycosylation present and other cellular molecules.

The finding that asparagine 631 was absent from the Lys-C digest is thought have arisen due to poor retention of this glycopeptide on the reversed-phase column. Due to the presence of the $Man_7GlcNAc_2$-M6P positional isomers and the low levels detected, the data indicating the presence of these oligosaccharides is of pure qualitative nature. The location of the phosphorylated oligosaccharides experimentally determined in the current study is in agreement with previous reports wherein a combination of endo-H digested MPR purified β-glucuronidase glycopeptides were identified using LC-MS/MS or when $^{32}PO_4$ incorporation was monitored[289, 292]. Both studies suggested the presence of phosphate on the *N*-glycans at asparagine 272, 420 or 631 of β-glucuronidase. However, the oligosaccharides carrying the M6P residue were not structurally characterized. The crystal structure of β-glucuronidase revealed the presence of a putative lysosomal targeting motif showing similarity to the lysosomal targeting motif of human Cathepsin D consisting of a $^{194}$KYPKGY$^{199}$ sequence and a β-hairpin between residues 179 to 204 in addition to a geometrically conserved lysine 530, Figure 29[290].

This lysosomal targeting motif is thought to form the basis of recognition of the acid hydrolase by GlcNAc-P-T, the initiator of M6P tagging. Although the oligosaccharides at all four *N*-glycosylation sites are extended away from the protein surface, it is thought that the close proximity of asparagine 173 to the putative lysosomal targeting motif shields it from the activity of GlcNAc-P-T. This may explain the absence of M6P at this *N*-glycosylation site and preferential phosphorylation of the glycans at asparagine 272, 420 and 631, Figure 30[290]. Additionally, the observation that *circa* 8.1% of all Asp 631 present displays glycans bearing M6P is corroborated by previous studies that also found the presence of M6P on the oligosaccharides at asparagine 631 to be a transient modification[289].

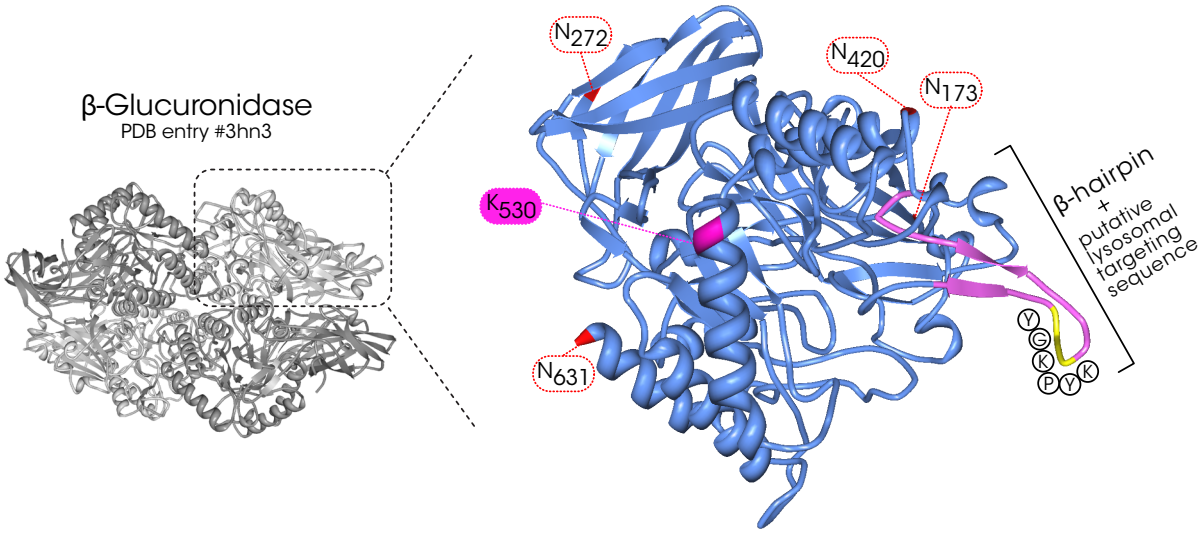



***Figure 29:*** *Molecular model of a β-glucuronidase monomer, Protein Data Bank entry number 3hn3, indicating the presence of each N-link glycosylation site as well as the β-hairpin containing the putative lysosomal targeting motif and lysine 530. Adapted from*[238].

In conclusion, a combination of UPLC using a 1.7 µm hydrophilic interaction phase and CE-LIF for the identification of M6P containing oligosaccharides on recombinant β-glucuronidase was employed. Traditional methods such as exoglycosidase digestion and offline anion exchange fractionation were not definitive for the presence of phosphate but were capable of identifying peaks in the *N*-glycan pool requiring further investigation. The additional use of mixed oxide affinity purification for the selective enrichment of phosphate containing oligosaccharides with subsequent alkaline phosphatase digestion was powerful, allowing for the identification of $Man_6GlcNAc_2$ and $Man_7GlcNAc_2$ glycans as the oligosaccharides displaying the M6P tag on β-glucuronidase. The combined use of UPLC-fluorescence and CE-LIF, due to the high degree of orthogonality in the resulting data sets, proved to be an important factor in identifying M6P on both $Man_6GlcNAc_2$ and $Man_7GlcNAc_2$. Considering the criticality of

M6P determination to ensure enzyme efficacy, the presented approach represents a significant analytical advance for the structural characterization of recombinant therapeutic enzymes used for lysosomal storage disorders treatment allowing for the determination of M6P at both the oligosaccharide and glycopeptide levels.

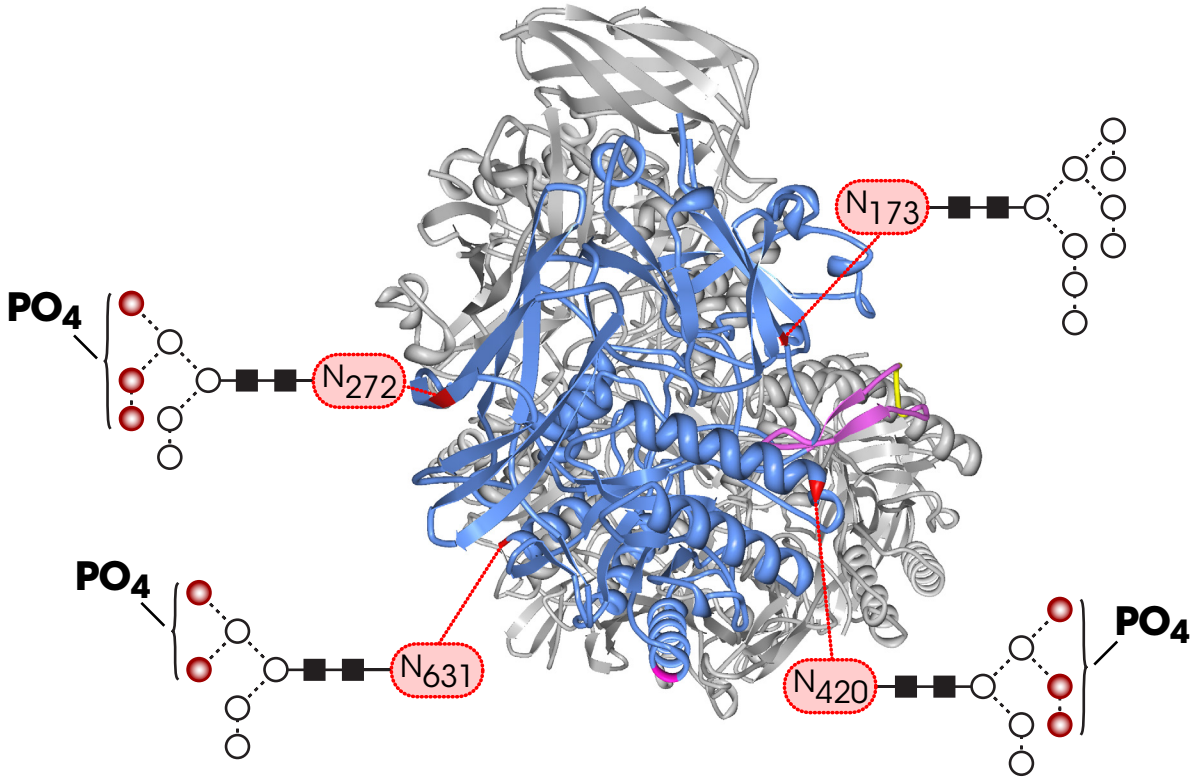



***Figure 30:*** *Molecular representation of the oligosaccharides present at each N-glycosylation site on a monomer of the β-glucuronidase tetramer. For $Man_7GlcNAc_2$-M6P the $Man_7GlcNAc_2$ D2 isomer is used for illustration, however, an equal probability exists for the presence of the $Man_7GlcNAc_2$ D1 or D3 isomers. A shaded mannose symbol represents a potential site of phosphate attachment. Adapted from*[238].

# CHAPTER 6

## Summary

Compared to HILIC-HPLC, significant improvements in efficiency, selectivity and analysis speed offered by UPLC profiling of fluorescently labeled serum *N*-glycans on a sub-2 μm HILIC phase enabled the identification of stomach adenocarcinoma associated alterations. Differential glycosylation present on four highly abundant proteins, IgG, haptoglobin, transferrin and α1-acid glycoprotein, isolated from pathologically staged cancer serum using either affinity purification or 2-DE, were screened for targeted stomach cancer progression markers.

The developed CE based method also enabled rapid and high efficiency separations of complex carbohydrate pools. Mere differences in linkage type or anomericity of isomeric oligosaccharides revealed differential migration, thus reflecting the separation principle based on discrepancies in hydrodynamic radii and associated molecular conformations. High precision induced by internal standards and GU normalization permits comparative studies of biomedical interest by CE. In a proof of concept study, sophisticated *N*-glycan features, such as core or antennary fucosylation, altered in inflammatory and malignant lung diseases, were successfully accentuated by CE-LIF. The standardized CE method not only revealed its potential for large scale studies, but also paves the way for a database enabling direct deduction of carbohydrate structural information from experimental CE data, similar to the HILIC-HPLC sources.

Both UPLC and CE-LIF revealed excellent but different selectivity for IgG *N*-glycans. Exhaustive glycan elucidation using an elaborate combination of weak anion exchange enrichment and exoglycosidase array digestion with subsequent profiling exceeded 48 hours. Combination of the data individually generated using each technique demonstrated that complete structural annotation was possible within a total analysis time of 20 minutes due to the advantageous orthogonality of the separation mechanisms. Parallel use of both analytical techniques offers a powerful platform for rapid, comprehensive and confident analysis of IgG *N*-glycans present on therapeutic antibodies or on antibodies of biomedical significance.

The orthogonal selectivity of UPLC and CE decisively contributed to characterizing mono- and *bis*-mannose-6-phosphate bearing oligosaccharides present on recombinant β-glucuronidase used for lysosomal storage disorder treatment. Exoglycosidase digestion and anion exchange fractionation were not definitive for identifying M6P decorated glycans. Mixed oxide affinity purification and alkaline phosphatase digestion, however, identified the glycan carriers of the M6P tag and glycoproteomic analysis revealed their respective site of attachment.

# CHAPTER 7

# Bibliography

## 7.1 Citations


1. El Rassi, Z., in *Carbohydrate Analysis*, ed. Z. El Rassi. Elsevier: Amsterdam, **1995**, vol. 58.
2. Harvey, D. J., Merry, A. H., Royle, L., Campbell, M. P., Dwek, R. A., Rudd, P. M., *Proteomics* **2009,** *9*. 3796-3801.
3. Apweiler, R., Hermjakob, H., Sharon, N., *Biochim. Biophys. Acta, Gen. Subj.* **1999,** *1473*. 4-8.
4. Varki, N. M., Varki, A., *Laboratory Investigation* **2007,** *87*. 851-857.
5. Taylor, M. E., Drickamer, K., *Introduction to Glycobiology*. 2nd ed.; Oxford University Press: Oxford, **2006**.
6. Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., Etzler, M. E. Cold Spring Harbor Laboratory Press, 2nd edn., 2008.
7. Petrescu, A.-J., Wormald, M. R., Dwek, R. A., *Curr. Opin. Struct. Biol.* **2006,** *16*. 600-607.
8. Shental-Bechor, D., Levy, Y., *Curr. Opin. Struct. Biol.* **2009,** *19*. 524-533.
9. Huhn, C., Selman, M. H. J., Ruhaak, L. R., Deelder, A. M., Wuhrer, M., *Proteomics* **2009,** *9*. 882-913.
10. Hermeling, S., Crommelin, D. J. A., Schellekens, H., Jiskoot, W., *Pharm. Res.* **2004,** *21*. 897-903.
11. Molinari, M., Calanca, V., Galli, C., Lucca, P., Paganetti, P., *Science* **2003,** *299*. 1397-1400.
12. Nezlin, R., Ghetie, V., *Adv. Immunol.* **2004,** *82*. 155-215.
13. Raju, T. S., Scallon, B., *Biotechnol. Prog.* **2007,** *23*. 964-971.
14. Kobata, A., *Biochim. Biophys. Acta, Gen. Subj.* **2008,** *1780*. 472-478.
15. del Val, I. J., Kontoravdi, C., Nagy, J. M., *Biotechnol. Prog.* **2010,** *26*. 1505-1527.
16. Walsh, G., Jefferis, R., *Nat. Biotechnol.* **2006,** *24*. 1241-1252.
17. Jefferis, R., Lund, J., Pound, J. D., *Immunol. Rev.* **1998,** *163*. 59-76.
18. Lund, J., Takahashi, N., Pound, J. D., Goodall, M., Jefferis, R., *J. Immunol.* **1996,** *157*. 4963-4969.
19. Krapp, S., Mimura, Y., Jefferis, R., Huber, R., Sondermann, P., *J. Mol. Biol.* **2003,** *325*. 979-989.
20. Masuda, K., Yamaguchi, Y., Kato, K., Takahashi, N., Shimada, I., Arata, Y., *FEBS Lett.* **2000,** *473*. 349-357.
21. Ferrara, C., Stuart, F., Sondermann, P., Bruenker, P., Umana, P., *J. Biol. Chem.* **2006,** *281*. 5032-5036.
22. Nimmerjahn, F., Ravetch, J. V., *Nat. Rev. Immunol.* **2008,** *8*. 34-47.
23. Shinkawa, T., Nakamura, K., Yamane, N., Shoji-Hosaka, E., Kanda, Y., Sakurada, M., Uchida, K., Anazawa, H., Satoh, M., Yamasaki, M., Hanai, N., Shitara, K., *J. Biol. Chem.* **2003,** *278*. 3466-3473.

24. Nimmerjahn, F., Ravetch, J. V., *Curr. Opin. Immunol.* **2007,** *19*. 239-245.

25. Arnold, J. N., Dwek, R. A., Rudd, P. M., Sim, R. B., *Immunol. Lett.* **2006,** *106*. 103-110.

26. Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M., Dwek, R. A., *Annu. Rev. Immunol.* **2007,** *25*. 21-50.

27. Malhotra, R., Wormald, M. R., Rudd, P. M., Fischer, P. B., Dwek, R. A., Sim, R. B., *Nat. Med.* **1995,** *1*. 237-243.

28. Nimmerjahn, F., Anthony, R. M., Ravetch, J. V., *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104*. 8433-8437.

29. Kaneko, Y., Nimmerjahn, F., Ravetch, J. V., *Science* **2006,** *313*. 670-673.

30. Scallon, B. J., Tam, S. H., McCarthy, S. G., Cai, A. N., Raju, T. S., *Mol. Immunol.* **2007,** *44*. 1524-1534.

31. Langlois, M. R., Delanghe, J. R., *Clin. Chem.* **1996,** *42*. 1589-1600.

32. Nakano, M., Nakagawa, T., Ito, T., Kitada, T., Hijioka, T., Kasahara, A., Tajiri, M., Wada, Y., Taniguchi, N., Miyoshi, E., *Int. J. Cancer* **2008,** *122*. 2301-2309.

33. He, Z., Aristoteli, L. P., Kritharides, L., Garner, B., *Biochem. Biophys. Res. Commun.* **2006,** *343*. 496-503.

34. Kaartinen, V., Mononen, I., *Biochim. Biophys. Acta* **1988,** *953*. 345-352.

35. Turk, B., Turk, V., *J. Biol. Chem.* **2009,** *284*. 21783-21787.

36. Kornfeld, S., *FASEB Journal* **1987,** *1*. 462-468.

37. Bao, M., Booth, J. L., Elmendorf, B. J., Canfield, W. M., *J. Biol. Chem.* **1996,** *271*. 31437-31445.

38. Kornfeld, R., Bao, M., Brewer, K., Noll, C., Canfield, W. M., *J. Biol. Chem.* **1998,** *273*. 23203-23210.

39. Pohlmann, R., Wendland, M., Boeker, C., von Figura, K., *J. Biol. Chem.* **1995,** *270*. 27311-27318.

40. Roberts, D. L., Weix, D. J., Dahms, N. M., Kim, J.-J. P., *Cell* **1998,** *93*. 639-648.

41. Saftig, P., Klumperman, J., *Nat. Rev. Mol. Cell Biol.* **2009,** *10*. 623-635.

42. Sun, P., Sleat, D. E., Lecocq, M., Hayman, A. R., Jadot, M., Lobel, P., *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105*. 16590-16595.

43. Alavi, A., Axford, J. S., *Rheumatology* **2008,** *47*. 760-770.

44. Hanash, S. M., Pitteri, S. J., Faca, V. M., *Nature* **2008,** *452*. 571-579.

45. Dube, D. H., Bertozzi, C. R., *Nat. Rev. Drug Discovery* **2005,** *4*. 477-488.

46. Dennis, J. W., Laferte, S., Yagel, S., Breitman, M. L., *Cancer Cells* **1989,** *1*. 87-92.

47. Saldova, R., Royle, L., Radcliffe, C. M., Abd Hamid, U. M., Evans, R., Arnold, J. N., Banks, R. E., Hutson, R., Harvey, D. J., Antrobus, R., Petrescu, S. M., Dwek, R. A., Rudd, P. M., *Glycobiology* **2007,** *17*. 1344-1356.

48. Li, B., An, H. J., Kirmiz, C., Lebrilla, C. B., Lam, K. S., Miyamoto, S., *J. Proteome Res.* **2008,** *7*. 3776-3788.

49. An, H. J., Miyamoto, S., Lancaster, K. S., Kirmiz, C., Li, B., Lam, K. S., Leiserowitz, G. S., Lebrilla, C. B., *J. Proteome Res.* **2006,** *5*. 1626-1635.

50. Abd Hamid, U. M., Royle, L., Saldova, R., Radcliffe, C. M., Harvey, D. J., Storr, S. J., Pardo, M., Antrobus, R., Chapman, C. J., Zitzmann, N., Robertson, J. F., Dwek, R. A., Rudd, P. M., *Glycobiology* **2008,** *18*. 1105-1118.

51. Pierce, A., Saldova, R., Abd Hamid, U. M., Abrahams, J. L., McDermott, E. W., Evoy, D., Duffy, M. J., Rudd, P. M., *Glycobiology* **2010,** *20*. 1283-1288.

52. Alley, W. R., Jr., Novotny, M. V., *J. Proteome Res.* **2010,** *9*. 3062-3072.

53. Kyselova, Z., Mechref, Y., Kang, P., Goetz, J. A., Dobrolecki, L. E., Sledge, G. W., Schnaper, L., Hickey, R. J., Malkas, L. H., Novotny, M. V., *Clin. Chem.* **2008,** *54*. 1166-1175.

54. Alley, W. R., Jr., Madera, M., Mechref, Y., Novotny, M. V., *Anal. Chem.* **2010,** *82*. 5095-5106.

55. Storr, S. J., Royle, L., Chapman, C. J., Abd Hamid, U. M., Robertson, J. F., Murray, A., Dwek, R. A., Rudd, P. M., *Glycobiology* **2008,** *18*. 456-462.

56. Sarrats, A., Saldova, R., Pla, E., Fort, E., Harvey, D. J., Struwe, W. B., de Llorens, R., Rudd, P. M., Peracaula, R., *Proteomics: Clin. Appl.* **2010,** *4*. 432-448.

57. Zhao, J., Patwa, T. H., Qiu, W., Shedden, K., Hinderer, R., Misek, D. E., Anderson, M. A., Simeone, D. M., Lubman, D. M., *J. Proteome Res.* **2007,** *6*. 1864-1874.

58. Zhao, J., Qiu, W., Simeone, D. M., Lubman, D. M., *J. Proteome Res.* **2007,** *6*. 1126-1138.

59. Tabares, G., Radcliffe, C. M., Barrabes, S., Ramirez, M., Aleixandre, R. N., Hoesel, W., Dwek, R. A., Rudd, P. M., Peracaula, R., de Llorens, R., *Glycobiology* **2006,** *16*. 132-145.

60. Peracaula, R., Tabares, G., Royle, L., Harvey, D. J., Dwek, R. A., Rudd, P. M., de Llorens, R., *Glycobiology* **2003,** *13*. 457-470.

61. Kyselova, Z., Mechref, Y., Al Bataineh, M. M., Dobrolecki, L. E., Hickey, R. J., Vinson, J., Sweeney, C. J., Novotny, M. V., *J. Proteome Res.* **2007,** *6*. 1822-1832.

62. Norton, P. A., Comunale, M. A., Krakover, J., Rodemich, L., Pirog, N., D'Amelio, A., Philip, R., Mehta, A. S., Block, T. M., *J. Cell. Biochem.* **2008,** *104*. 136-149.

63. Wang, M., Long, R. E., Comunale, M. A., Junaidi, O., Marrero, J., Di Bisceglie, A. M., Block, T. M., Mehta, A. S., *Cancer Epidemiol., Biomarkers Prev.* **2009,** *18*. 1914-1921.

64. Goldman, R., Ressom, H. W., Varghese, R. S., Goldman, L., Bascug, G., Loffredo, C. A., Abdel-Hamid, M., Gouda, I., Ezzat, S., Kyselova, Z., Mechref, Y., Novotny, M. V., *Clin. Cancer Res.* **2009,** *15*. 1808-1813.

65. Hongsachart, P., Huang-Liu, R., Sinchaikul, S., Pan, F.-M., Phutrakul, S., Chuang, Y.-M., Yu, C.-J., Chen, S.-T., *Electrophoresis* **2009,** *30*. 1206-1220.

66. Kim, Y. J., Varki, A., *Glycoconjugate J.* **1997,** *14*. 569-576.

67. Seales, E. C., Jurado, G. A., Brunson, B. A., Wakefield, J. K., Frost, A. R., Bellis, S. L., *Cancer Res.* **2005,** *65*. 4645-4652.

68. Bones, J., Byrne, J. C., O'Donoghue, N., McManus, C., Scaife, C., Boissin, H., Nastase, A., Rudd, P. M., *J. Proteome Res.* **2011,** *10*. 1246-1265.

69. Bones, J., Mittermayr, S., O'Donoghue, N., Guttman, A., Rudd, P. M., *Anal. Chem.* **2010,** *82*. 10208-10215.

70. van de Geijn, F. E., Wuhrer, M., Selman, M. H. J., Willemsen, S. P., de Man, Y. A., Deelder, A. M., Hazes, J. M. W., Dolhain, R. J. E. M., *Arthritis Res. Ther.* **2009,** *11*. No pp given.

71. Wuhrer, M., Stam, J. C., van de Geijn, F. E., Koeleman, C. A. M., Verrips, C. T., Dolhain, R. J. E. M., Hokke, C. H., Deelder, A. M., *Proteomics* **2007,** *7*. 4070-4081.

72. Angelini, C., Semplicini, C., *Curr. Opin. Pharmacol.* **2010,** *10*. 338-345.

73. Brady, R. O., *Annu. Rev. Med.* **2006,** *57*. 283-296.

74. Barton, N. W., Brady, R. O., Dambrosia, J. M., Di Bisceglie, A. M., Doppelt, S. H., Hill, S. C., Mankin, H. J., Murray, G. J., Parker, R. I., Argoff, C. E., et al., *N. Engl. J. Med.* **1991,** *324*. 1464-1470.

75. Cox, T. M., *Biol.: Targets Ther.* **2010,** *4*. 299-313.

76. Parenti, G., *EMBO Mol. Med.* **2009,** *1*. 268-279.

77. Zhu, Y., Li, X., Kyazike, J., Zhou, Q., Thurberg, B. L., Raben, N., Mattaliano, R. J., Cheng, S. H., *J. Biol. Chem.* **2004,** *279*. 50336-50341.

78. Zhu, Y., Li, X., McVie-Wylie, A., Jiang, C., Thurberg, B. L., Raben, N., Mattaliano, R. J., Cheng, S. H., *Biochem. J.* **2005,** *389*. 619-628.

79. Taylor Allen, D., Hancock William, S., Hincapie, M., Taniguchi, N., Hanash Samir, M., *Genome Med.* **2009,** *1*. 57.

80. Novotny, M. V., Soini, H. A., Mechref, Y., *J. Chromatogr., B Anal. Technol. Biomed. Life Sci.* **2008,** *866*. 26-47.

81. Kang, P., Mechref, Y., Novotny, M. V., *Rapid Commun. Mass Spectrom.* **2008,** *22*. 721-734.

82. Domann Paula, J., Pardos-Pardos Ana, C., Fernandes Daryl, L., Spencer Daniel Ian, R., Radcliffe Catherine, M., Royle, L., Dwek Raymond, A., Rudd Pauline, M., *Proteomics* **2007,** *7 Suppl 1*. 70-76.

83. Ruhaak, L. R., Huhn, C., Waterreus, W.-J., de Boer, A. R., Neususs, C., Hokke, C. H., Deelder, A. M., Wuhrer, M., *Anal. Chem.* **2008,** *80*. 6119-6126.

84. Alpert, A. J., *J. Chromatogr.* **1990,** *499*. 177-196.

85. Alpert, A. J., Shukla, M., Shukla, A. K., Zieske, L. R., Yuen, S. W., Ferguson, M. A. J., Mehlert, A., Pauly, M., Orlando, R., *J. Chromatogr., A* **1994,** *676*. 191-202.

86. Campbell, M. P., Royle, L., Radcliffe, C. M., Dwek, R. A., Rudd, P. M., *Bioinformatics* **2008,** *24*. 1214-1216.

87. Royle, L., Campbell, M. P., Radcliffe, C. M., White, D. M., Harvey, D. J., Abrahams, J. L., Kim, Y.-G., Henry, G. W., Shadick, N. A., Weinblatt, M. E., Lee, D. M., Rudd, P. M., Dwek, R. A., *Anal. Biochem.* **2008,** *376*. 1-12.

88. Wilson, I. D., Nicholson, J. K., Castro-Perez, J., Granger, J. H., Johnson, K. A., Smith, B. W., Plumb, R. S., *J. Proteome Res.* **2005,** *4*. 591-598.

89. Neue, U. D., Kele, M., Bunner, B., Kromidas, A., Dourdeville, T., Mazzeo, J. R., Grumbach, E. S., Serpa, S., Wheat, T. E., Hong, P., Gilar, M., *Adv. Chromatogr.* **2010,** *48*. 99-143.

90. Guillarme, D., Schappler, J., Rudaz, S., Veuthey, J.-L., *TrAC, Trends Anal. Chem.* **2010,** *29*. 15-27.

91. Sandra, K., Moshir, M., D'Hondt, F., Verleysen, K., Kas, K., Sandra, P., *J. Chromatogr., B Anal. Technol. Biomed. Life Sci.* **2008,** *866*. 48-63.

92. Rossi, C., Calton, L., Hammond, G., Brown, H. A., Wallace, A. M., Sacchetta, P., Morris, M., *Clin. Chim. Acta* **2010,** *411*. 222-228.

93. Zelena, E., Dunn, W. B., Broadhurst, D., Francis-McIntyre, S., Carroll, K. M., Begley, P., O'Hagan, S., Knowles, J. D., Halsall, A., Wilson, I. D., Kell, D. B., *Anal. Chem.* **2009,** *81*. 1357-1364.

94. Want, E. J., Coen, M., Masson, P., Keun, H. C., Pearce, J. T. M., Reily, M. D., Robertson, D. G., Rohde, C. M., Holmes, E., Lindon, J. C., Plumb, R. S., Nicholson, J. K., *Anal. Chem.* **2010,** *82*. 5282-5289.

95. Zhang, J., Yan, L., Chen, W., Lin, L., Song, X., Yan, X., Hang, W., Huang, B., *Anal. Chim. Acta* **2009,** *650*. 16-22.

96. Chen, J., Wang, W., Lv, S., Yin, P., Zhao, X., Lu, X., Zhang, F., Xu, G., *Anal. Chim. Acta* **2009,** *650*. 3-9.

97. Sandra, K., Pereira, A. d. S., Vanhoenacker, G., David, F., Sandra, P., *J. Chromatogr., A* **2010,** *1217*. 4087-4099.

98. Lankinen, M., Schwab, U., Erkkila, A., Seppanen-Laakso, T., Hannila, M.-L., Mussalo, H., Lehto, S., Uusitupa, M., Gylling, H., Oresic, M., *PloS one* **2009,** *4*. e5258.

99. Lengqvist, J., Andrade, J., Yang, Y., Alvelius, G., Lewensohn, R., Lehtioe, J., *J. Chromatogr., B Anal. Technol. Biomed. Life Sci.* **2009,** *877*. 1306-1316.

100. Bordeerat, N. K., Georgieva, N. I., Klapper, D. G., Collins, L. B., Cross, T. J., Borchers, C. H., Swenberg, J. A., Boysen, G., *Proteomics* **2009,** *9*. 3939-3944.

101. Kasper, D. C., Herman, J., De Jesus, V. R., Mechtler, T. P., Metz, T. F., Shushan, B., *Rapid Commun. Mass Spectrom.* **2010,** *24*. 986-994.

102. Kurihara, T., Min, J. Z., Hirata, A., Toyo'oka, T., Inagaki, S., *Biomed. Chromatogr.* **2009,** *23*. 516-523.

103. Min, J. Z., Toyo'oka, T., Kurihara, T., Kato, M., Fukushima, T., Inagaki, S., *Biomed. Chromatogr.* **2007,** *21*. 852-860.

104. Jones, C. J., Membreno, N., Larive, C. K., *J. Chromatogr., A* **2010,** *1217*. 479-488.

105. Ahn, J., Bones, J., Yu, Y. Q., Rudd, P. M., Gilar, M., *J. Chromatogr., B Anal. Technol. Biomed. Life Sci.* **2010,** *878*. 403-408.

106. Melmer, M., Stangler, T., Premstaller, A., Lindner, W., *J. Chromatogr., A* **2011,** *1218*. 118-123.

107. Tiselius, A., *Trans. Faraday Soc.* **1937,** *33*. 524-531.

108. Andrews, A. T., *Electrophoresis. Theory, techniques, and biochemical and clinical applications*. 2 ed.; Clarendon Press: Oxford, **1986**.

109. Raju, T. S., *Anal. Biochem.* **2000,** *283*. 125-132.

110. Kuhn, R., Hoffstetter-Kuhn, S., *Capillary Electrophoresis: Principles and Practice*. Springer-Verlag: Berlin, New York, Heidelberg, **1993**.

111. Jorgenson, J. W., Lukacs, K. D., *Science* **1983,** *222*. 266-72.

112. Chen, F.-T. A., Evangelista, R. A., *Anal. Biochem.* **1995,** *230*. 273-280.

113. Huhn, C., Ramautar, R., Wuhrer, M., Somsen, G. W., *Anal. Bioanal. Chem.* **2010,** *396*. 297-314.

114. Chiesa, C., Horvath, C., *J. Chromatogr.* **1993,** *645*. 337-352.

115. Camilleri, P., Harland, G. B., Okafo, G., *Anal. Biochem.* **1995,** *230*. 115-122.

116. Klockow, A., Amado, R., Widmer, H. M., Paulus, A., *Electrophoresis* **1996,** *17*. 110-119.

117. Charlwood, J., Birrell, H., Organ, A., Camilleri, P., *Rapid Commun. Mass Spectrom.* **1999,** *13*. 716-723.

118. Harland, G. B., Okafo, G., Matejtschuk, P., Sellick, I. C., Chapman, G. E., Camilleri, P., *Electrophoresis* **1996,** *17*. 406-411.

119. Honda, S., Okeda, J., Iwanaga, H., Kawakami, S., Taga, A., Suzuki, S., Imai, K., *Anal. Biochem.* **2000,** *286*. 99-111.

120. Tran, N. T., Taverna, M., Deschamps, F. S., Morin, P., Ferrier, D., *Electrophoresis* **1998,** *19*. 2630-2638.

121. Zhong, H., El Rassi, Z., *J. Sep. Sci.* **2009,** *32*. 10-20.

122. Guttman, A., *Nature* **1996,** *380*. 461-462.

123. Nakano, M., Higo, D., Arai, E., Nakagawa, T., Kakehi, K., Taniguchi, N., Kondo, A., *Glycobiology* **2009,** *19*. 135-143.

124. Oyama, T., Yodohsi, M., Yamane, A., Kakehi, K., Hayakawa, T., Suzuki, S., *J. Chromatogr., B Analyt. Technol. Biomed. Life Sci.* **2011,** *879*. 2928-2934.

125. Anumula, K. R., *Anal. Biochem.* **2006,** *350*. 1-23.

126. Ruhaak, L. R., Zauner, G., Huhn, C., Bruggink, C., Deelder, A. M., Wuhrer, M., *Anal. Bioanal. Chem.* **2010,** *397*. 3457-3481.

127. Harvey, D. J., *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2011,** *879*. 1196-1225.

128. Hara, S., Yamaguchi, M., Takemori, Y., Nakamura, M., Ohkura, Y., *J. Chromatogr., Biomed. Appl.* **1986,** *377*. 111-19.

129. Zanetta, J.-P., Pons, A., Iwersen, M., Mariller, C., Leroy, Y., Timmerman, P., Schauer, R., *Glycobiology* **2001,** *11*. 663-676.

130. Honda, S., *J. Chromatogr., A* **1996,** *720*. 337-51.

131. Pioch, M., Bunz, S.-C., Neususess, C., *Electrophoresis* **2012,** *33*. 1517-1530.

132. Que, A. H., Novotny, M. V., *Anal. Bioanal. Chem.* **2003,** *375*. 599-608.

133. Yoshino, K.-i., Takao, T., Murata, H., Shimonishi, Y., *Anal. Chem.* **1995,** *67*. 4028-31.

134. Krenkova, J., Foret, F., *Proteomics* **2012,** *12*. 2978-2990.

135. Mechref, Y., Novotny, M. V., *Mass Spectrom. Rev.* **2009,** *28*. 207-222.

136. Zhao, S. S., Zhong, X., Tie, C., Chen, D. D. Y., *Proteomics* **2012,** *12*. 2991-3012.

137. Bindila, L., Peter-Katalinic, J., *Mass Spectrom. Rev.* **2009,** *28*. 223-253.

138. Dang, F., Kakehi, K., Nakajima, K., Shinohara, Y., Ishikawa, M., Kaji, N., Tokeshi, M., Baba, Y., *J. Chromatogr., A* **2006,** *1109*. 138-143.

139. Primack, J., Flynn, G. C., Pan, H., *Electrophoresis* **2011,** *32*. 1129-1132.

140. Vanderschaeghe, D., Szekrenyes, A., Wenz, C., Gassmann, M., Naik, N., Bynum, M., Yin, H., Delanghe, J., Guttman, A., Callewaert, N., *Anal. Chem.* **2010,** *82*. 7408-7415.

141. Mitra, I., Zhuang, Z., Zhang, Y., Yu, C.-Y., Hammoud, Z. T., Tang, H., Mechref, Y., Jacobson, S. C., *Anal. Chem.* **2012,** *84*. 3621-3627.

142. Zhuang, Z., Starkey, J. A., Mechref, Y., Novotny, M. V., Jacobson, S. C., *Anal. Chem.* **2007,** *79*. 7170-7175.

143. Zhuang, Z.-X., Mitra, I., Hussein, A., Novotny, M. V., Mechref, Y., Jacobson, S. C., *Electrophoresis* **2011,** *32*. 246-253.

144. Fang, M., Zhao, Y.-P., Zhou, F.-G., Lu, L.-G., Qi, P., Wang, H., Zhou, K., Sun, S.-H., Chen, C.-Y., Gao, C.-F., *Int. J. Cancer* **2010,** *127*. 148-159.

145. Vanderschaeghe, D., Laroy, W., Sablon, E., Halfon, P., Van Hecke, A., Delanghe, J., Callewaert, N., *Mol. Cell. Proteomics* **2009,** *8*. 986-994.

146. Ruhaak, L. R., Hennig, R., Huhn, C., Borowiak, M., Dolhain, R., Deelder, A. M., Rapp, E., Wuhrer, M., *J. Proteome Res.* **2010,** *9*. 6655-6664.

147. Hermentin, P., Doenges, R., Witzel, R., Hokke, C. H., Vliegenthart, J. F. G., Kamerling, J. P., Conradt, H. S., Nimtz, M., Brazel, D., *Anal. Biochem.* **1994,** *221*. 29-41.

148. Szabo, Z., Guttman, A., Bones, J., Karger, B. L., *Anal. Chem.* **2011,** *83*. 5329-5336.

149. Taverna, M., Baillet, A., Baylocq-Ferrier, D., *Chromatographia* **1993,** *37*. 415-22.

150. Guttman, A., Pritchett, T., *Electrophoresis* **1995,** *16*. 1906-11.

151. Wolff, M. W., Murhammer, D. W., Linhardt, R. J., *Prep. Biochem. Biotechnol.* **1999,** *29*. 1-21.

152. Schwarzer, J., Rapp, E., Reichl, U., *Electrophoresis* **2008,** *29*. 4203-4214.

153. Suzuki, S., Honda, S., *TrAC, Trends Anal. Chem.* **1995,** *14*. 279-88.

154. Nashabeh, W., El Rassi, Z., *J. Chromatogr.* **1992,** *600*. 279-87.

155. Laroy, W., Contreras, R., Callewaert, N., *Nat. Protoc.* **2006,** *1*. 397-405.

156. Khandurina, J., Olson, N. A., Anderson, A. A., Gray, K. A., Guttman, A., *Electrophoresis* **2004,** *25*. 3117-3121.

157. Kovats, E., *Helv. Chim. Acta* **1958,** *41*. 1915-32.

158. Ma, S., Nashabeh, W., *Anal. Chem.* **1999,** *71*. 5185-5192.

159. Szabo, Z., Guttman, A., Bones, J., Shand, R. L., Meh, D., Karger, B. L., *Mol. Pharmaceutics* **2012,** *9*. 1612-1619.

160. Mikkers, F. E. P., Everaerts, F. M., Verheggen, T. P. E. M., *J. Chromatogr.* **1979,** *169*. 1-10.

161. Guttman, A., *Electrophoresis* **1997,** *18*. 1136-1141.

162. Callewaert, N., Van Vlierberghe, H., Van Hecke, A., Laroy, W., Delanghe, J., Contreras, R., *Nat. Med.* **2004,** *10*. 429-434.

163. Dang, F., Kakehi, K., Cheng, J., Tabata, O., Kurokawa, M., Nakajima, K., Ishikawa, M., Baba, Y., *Anal. Chem.* **2006,** *78*. 1452-1458.

164. Luo, R., Archer-Hartmann, S. A., Holland, L. A., *Anal. Chem.* **2010,** *82*. 1228-1233.

165. Zhao, Y.-P., Ruan, C.-P., Wang, H., Hu, Z.-Q., Fang, M., Gu, X., Ji, J., Zhao, J.-Y., Gao, C.-F., *Cancer* **2012,** *118*. 639-650.

166. Adamczyk, B., Tharmalingam, T., Rudd, P. M., *Biochim. Biophys. Acta, Gen. Subj.* **2012,** *1820*. 1347-1353.

167. Varki, A., *Trends Mol. Med.* **2008,** *14*. 351-360.

168. O'Neill, R. A., *J. Chromatogr., A* **1996,** *720*. 201-15.

169. Frisch, E., Kaup, M., Egerer, K., Weimann, A., Tauber, R., Berger, M., Blanchard, V., *Electrophoresis* **2011,** *32*. 3510-3515.

170. Fenaille, F., Le Mignon, M., Groseil, C., Ramon, C., Riande, S., Siret, L., Bihoreau, N., *Glycobiology* **2007,** *17*. 932-944.

171. Kakehi, K., Funakubo, T., Suzuki, S., Oda, Y., Kitada, Y., *J. Chromatogr., A* **1999,** *863*. 205-218.

172. Yagi, Y., Yamamoto, S., Kakehi, K., Hayakawa, T., Ohyama, Y., Suzuki, S., *Electrophoresis* **2011,** *32*. 2979-2985.

173. Archer-Hartmann, S. A., Crihfield, C. L., Holland, L. A., *Electrophoresis* **2011,** *32*. 3491-3498.

174. Archer-Hartmann, S. A., Sargent, L. M., Lowry, D. T., Holland, L. A., *Anal. Chem.* **2011,** *83*. 2740-2747.

175. Hong, M., Cassely, A., Mechref, Y., Novotny, M. V., *J. Chromatogr., B: Biomed. Sci. Appl.* **2001,** *752*. 207-216.

176. Nakajima, K., Oda, Y., Kinoshita, M., Kakehi, K., *J. Proteome Res.* **2003,** *2*. 81-88.

177. Nakajima, K., Kinoshita, M., Oda, Y., Masuko, T., Kaku, H., Shibuya, N., Kakehi, K., *Glycobiology* **2004,** *14*. 793-804.

178. Dang, F., Maeda, E., Osafune, T., Nakajima, K., Kakehi, K., Ishikawa, M., Baba, Y., *Anal. Chem.* **2009,** *81*. 10055-10060.

179. Ishizuka, A., Hashimto, Y., Naka, R., Kinoshita, M., Kakehi, K., Seino, J., Funakoshi, Y., Suzuki, T., Kameyama, A., Narimatsu, H., *Biochem. J.* **2008,** *413*. 227-237.

180. Bones, J., McLoughlin, N., Hilliard, M., Wynne, K., Karger, B. L., Rudd, P. M., *Anal. Chem.* **2011,** *83*. 4154-4162.

181. Hermentin, P., Witzel, R., Doenges, R., Bauer, R., Haupt, H., Patel, T., Parekh, R. B., Brazel, D., *Anal. Biochem.* **1992,** *206*. 419-429.

182. Taverna, M., Baillet, A., Biou, D., Schluter, M., Werner, R., Ferrier, D., *Electrophoresis* **1992,** *13*. 359-366.

183. Liu, Y., Salas-Solano, O., Gennaro, L. A., *Anal. Chem.* **2009,** *81*. 6823-6829.

184. Suzuki, S., Tanaka, R., Takada, K., Inoue, N., Yashima, Y., Honda, A., Honda, S., *J. Chromatogr., A* **2001,** *910*. 319-329.

185. Kamoda, S., Ishikawa, R., Kakehi, K., *J. Chromatogr., A* **2006,** *1133*. 332-339.

186. Liu, X.-E., Desmyter, L., Gao, C.-F., Laroy, W., Dewaele, S., Vanhooren, V., Wang, L., Zhuang, H., Callewaert, N., Libert, C., Contreras, R., Chen, C., *Hepatology* **2007,** *46*. 1426-1435.

187. Quintero, O., Montesino, R., Cremata, J. A., *Anal. Biochem.* **1998,** *256*. 23-32.

188. Maeda, E., Kita, S., Kinoshita, M., Urakami, K., Hayakawa, T., Kakehi, K., *Anal. Chem.* **2012,** *84*. 2373-2379.

189. Shi, X., Zaia, J., *J. Biol. Chem.* **2009,** *284*. 11806-11814.

190. Honda, S., Makino, A., Suzuki, S., Kakehi, K., *Anal. Biochem.* **1990,** *191*. 228-234.

191. Chen, X., Flynn, G. C., *Anal. Biochem.* **2007,** *370*. 147-161.

192. Prater, B. D., Connelly, H. M., Qin, Q., Cockrill, S. L., *Anal. Biochem.* **2009,** *385*. 69-79.

193. Okafo, G., Burrow, L., Carr, S. A., Roberts, G. D., Johnson, W., Camilleri, P., *Anal. Chem.* **1996,** *68*. 4424-4430.

194. Guttman, A., Sperling, E., Mazsaroff, I., *J. Liq. Chromatogr. Relat. Technol.* **1996,** *19*. 1539-1549.

195. Wolff, M. W., Bazin, H. G., Lindhardt, R. J., *Biotechnol. Tech.* **1999,** *13*. 797-801.

196. An, H. J., Franz, A. H., Lebrilla, C. B., *J. Chromatogr., A* **2003,** *1004*. 121-129.

197. Suzuki, H., Mueller, O., Guttman, A., Karger, B. L., *Anal. Chem.* **1997,** *69*. 4554-4559.

198. Tomiya, N., Awaya, J., Kurono, M., Endo, S., Arata, Y., Takahashi, N., *Anal. Biochem.* **1988,** *171*. 73-90.

199. Suzuki, S., Kakehi, K., Honda, S., *Anal. Biochem.* **1992,** *205*. 227-236.

200. Zieske, L. R., Fu, D., Khan, S. H., O'Neill, R. A., *J. Chromatogr., A* **1996,** *720*. 395-407.

201. Briggs, J. B., Keck, R. G., Ma, S., Lau, W., Jones, A. J. S., *Anal. Biochem.* **2009,** *389*. 40-51.

202. Zaia, J., *Chem. Biol.* **2008,** *15*. 881-892.

203. Domon, B., Costello, C. E., *Glycoconjugate J.* **1988,** *5*. 397-409.

204. Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., Haslam, S. M., *J. Proteome Res.* **2008,** *7*. 1650-1659.

205. Maass, K., Ranzinger, R., Geyer, H., von der Lieth, C.-W., Geyer, R., *Proteomics* **2007,** *7*. 4435-4444.

206. Gennaro, L. A., Salas-Solano, O., Ma, S., *Anal. Biochem.* **2006,** *355*. 249-258.

207. Jayo, R. G., Li, J., Chen, D. D. Y., *Anal. Chem.* **2012,** *84*. 8756-8762.

208. Albrecht, S., Schols, H. A., van den Heuvel, E. G. H. M., Voragen, A. G. J., Gruppen, H., *Electrophoresis* **2010,** *31*. 1264-1273.

209. Gennaro, L. A., Salas-Solano, O., *Anal. Chem.* **2008,** *80*. 3838-3845.

210. Huhn, C., Ruhaak, L. R., Mannhardt, J., Wuhrer, M., Neususs, C., Deelder, A. M., Meyer, H., *Electrophoresis* **2012,** *33*. 563-566.

211. Que, A. H., Mechref, Y., Huang, Y., Taraszka, J. A., Clemmer, D. E., Novotny, M. V., *Anal. Chem.* **2003,** *75*. 1684-1690.

212. Sandra, K., Dolashka-Angelova, P., Devreese, B., Van Beeumen, J., *Glycobiology* **2007,** *17*. 141-156.

213. Gennaro, L. A., Delaney, J., Vouros, P., Harvey, D. J., Domon, B., *Rapid Commun. Mass Spectrom.* **2002,** *16*. 192-200.

214. Sandra, K., Van Beeumen, J., Stals, I., Sandra, P., Claeyssens, M., Devreese, B., *Anal. Chem.* **2004,** *76*. 5878-5886.

215. Harvey, D. J., *J. Am. Soc. Mass Spectrom.* **2000,** *11*. 900-915.

216. Balaguer, E., Demelbauer, U., Pelzing, M., Sanz-Nebot, V., Barbosa, J., Neususs, C., *Electrophoresis* **2006,** *27*. 2638-2650.

217. Balaguer, E., Neususs, C., *Anal. Chem.* **2006,** *78*. 5384-5393.

218. Ongay, S., Martin-Alvarez, P. J., Neususs, C., de Frutos, M., *Electrophoresis* **2010,** *31*. 3314-3325.

219. Ongay, S., Neususs, C., *Anal. Bioanal. Chem.* **2010,** *398*. 845-855.

220. Leymarie, N., Zaia, J., *Anal. Chem.* **2012,** *84*. 3040-3048.

221. Wheeler, S. F., Domann, P., Harvey, D. J., *Rapid Commun. Mass Spectrom.* **2009,** *23*. 303-312.

222. Ashline, D., Singh, S., Hanneman, A., Reinhold, V., *Anal. Chem.* **2005,** *77*. 6250-6262.

223. Tegeler, T. J., Mechref, Y., Boraas, K., Reilly, J. P., Novotny, M. V., *Anal. Chem.* **2004,** *76*. 6698-6706.

224. Kang, P., Mechref, Y., Klouckova, I., Novotny, M. V., *Rapid Commun. Mass Spectrom.* **2005,** *19*. 3421-3428.

225. Royle, L., Radcliffe, C. M., Dwek, R. A., Rudd, P. M., *Methods Mol. Biol.* **2006,** *347*. 125-143.

226. Olajos, M., Hajos, P., Bonn, G. K., Guttman, A., *Anal. Chem.* **2008,** *80*. 4241-4246.

227. Kirschner, K. N., Yongye, A. B., Tschampel, S. M., Gonzalez-Outeirino, J., Daniels, C. R., Foley, L. B., Woods, R. J., *J. Comput. Chem.* **2008,** *29*. 622-655.

228. Fournier, T., Medjoubi-N, N., Porquet, D., *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **2000,** *1482*. 157-171.

229. Mittermayr, S., Guttman, A., *Electrophoresis* **2012,** *33*. 1000-1007.

230. Mittermayr, S., Olajos, M., Chován, T., Bonn, G. K., Guttman, A., *Trends Anal. Chem.* **2008,** *27*. 407-417.

231. Kasicka, V., *Electrophoresis* **1999,** *20*. 3084-3105.

232. Varadi, C., Mittermayr, S., Szekrenyes, A., Kadas, J., Takacs, L., Kurucz, I., Guttman, A., *Electrophoresis* **2013,** *Accepted*.

233. Mittermayr, S., Bones, J., Doherty, M., Guttman, A., Rudd, P. M., *J. Proteome Res.* **2011,** *10*. 3820-3829.

234. Andre, S., Kozar, T., Kojima, S., Unverzagt, C., Gabius, H.-J., *Biol. Chem.* **2009,** *390*. 557-565.

235. Andre, S., Unverzagt, C., Kojima, S., Frank, M., Seifert, J., Fink, C., Kayser, K., von der Lieth, C.-W., Gabius, H.-J., *Eur. J. Biochem.* **2004,** *271*. 118-134.

236. Stadlmann, J., Pabst, M., Altmann, F., *J. Clin. Immunol.* **2010,** *30*. S15-S19.

237. Stadlmann, J., Weber, A., Pabst, M., Anderle, H., Kunert, R., Ehrlich, H. J., Schwarz, H. P., Altmann, F., *Proteomics* **2009,** *9*. 4143-4153.

238. Bones, J., Mittermayr, S., McLoughlin, N., Hilliard, M., Wynne, K., Johnson, G. R., Grubb, J. H., Sly, W. S., Rudd, P. M., *Anal. Chem.* **2011,** *83*. 5344-5352.

239. Knezevic, A., Polasek, O., Gornik, O., Rudan, I., Campbell, H., Hayward, C., Wright, A., Kolcic, I., O'Donoghue, N., Bones, J., Rudd, P. M., Lauc, G., *J. Proteome Res.* **2009,** *8*. 694-701.

240. Gornik, O., Wagner, J., Pucic, M., Knezevic, A., Redzic, I., Lauc, G., *Glycobiology* **2009,** *19*. 1547-1553.

241. Knezevic, A., Gornik, O., Polasek, O., Pucic, M., Redzic, I., Novokmet, M., Rudd, P. M., Wright, A. F., Campbell, H., Rudan, I., Lauc, G., *Glycobiology* **2010,** *20*. 959-969.

242. Pucic, M., Pinto, S., Novokmet, M., Knezevic, A., Gornik, O., Polasek, O., Vlahovicek, K., Wang, W., Rudd, P. M., Wright, A. F., Campbell, H., Rudan, I., Lauc, G., *Glycobiology* **2010,** *20*. 970-975.

243. Gu, J., Sato, Y., Kariya, Y., Isaji, T., Taniguchi, N., Fukuda, T., *J. Proteome Res.* **2009,** *8*. 431-435.

244. Takahashi, M., Kuroki, Y., Ohtsubo, K., Taniguchi, N., *Carbohydr. Res.* **2009,** *344*. 1387-1390.

245. Dennis, J. W., Granovsky, M., Warren, C. E., *BioEssays* **1999,** *21*. 412-421.

246. Dennis, J. W., Granovsky, M., Warren, C. E., *Biochim. Biophys. Acta, Gen. Subj.* **1999,** *1473*. 21-34.

247. Dwek, R. A., Lellouch, A. C., Wormald, M. R., *J. Anat.* **1995,** *187*. 279-292.

248. Boone, M. A., Nymeyer, H., Striegel, A. M., *Carbohydr. Res.* **2008,** *343*. 132-138.

249. Boone, M. A., Striegel, A. M., *Macromolecules* **2006,** *39*. 4128-4131.

250. Naidoo, K. J., Kuttel, M., *J. Comput. Chem.* **2001,** *22*. 445-456.

251. Striegel, A. M., Boone, M. A., *Biopolymers* **2011,** *95*. 228-233.

252. Sakajiri, T., Kikuchi, T., Simon, I., Uchida, K., Yamamura, T., Ishii, T., Yajima, H., *J. Mol. Struct.* **2006,** *764*. 133-140.

253. Vetter, D., Thorn, W., *Starch - Staerke* **1992,** *44*. 271-274.

254. Pereira, C. S., Kony, D., Baron, R., Muller, M., van Gunsteren, W. F., Hunenberger, P. H., *Biophys. J.* **2006,** *90*. 4337-4344.

255. Best, R. B., Jackson, G. E., Naidoo, K. J., *J. Phys. Chem. B* **2001,** *105*. 4742-4751.

256. Park, S.-Y., Yoon, S.-J., Jeong, Y.-T., Kim, J.-M., Kim, J.-Y., Bernert, B., Ullman, T., Itzkowitz, S. H., Kim, J.-H., Hakomori, S.-i., *Int. J. Cancer* **2010,** *126*. 142-155.

257. Lin, Z., Simeone, D. M., Anderson, M. A., Brand, R. E., Xie, X., Shedden, K. A., Ruffin, M. T., Lubman, D. M., *J. Proteome Res.* **2011,** *10*. 2602-2611.

258. Fujimura, T., Shinohara, Y., Tissot, B., Pang, P.-C., Kurogochi, M., Saito, S., Arai, Y., Sadilek, M., Murayama, K., Dell, A., Nishimura, S.-I., Hakomori, S.-I., *Int. J. Cancer* **2007,** *122*. 39-49.

259. Ruhaak, L. R., Uh, H.-W., Beekman, M., Hokke, C. H., Westendorp, R. G. J., Houwing-Duistermaat, J., Wuhrer, M., Deelder, A. M., Slagboom, P. E., *J. Proteome Res.* **2011,** *10*. 1667-1674.

260. Domann, P. J., Pardos-Pardos, A. C., Fernandes, D. L., Spencer, D. I. R., Radcliffe, C. M., Royle, L., Dwek, R. A., Rudd, P. M., *Practical Proteomics* **2007,** *2*. 70-76.

261. Marino, K., Bones, J., Kattla, J. J., Rudd, P. M., *Nat. Chem. Biol.* **2010,** *6*. 713-723.

262. Vanderschaeghe, D., Festjens, N., Delanghe, J., Callewaert, N., *Biol. Chem.* **2010,** *391*. 149-161.

263. Stefansson, M., Novotny, M., *Anal. Chem.* **1994,** *66*. 1134-40.

264. Lee, J., Park, S.-H., Stanley, P., *Glycoconjugate J.* **2003,** *19*. 211-219.

265. Stanley, P., *Biochim. Biophys. Acta, Gen. Subj.* **2002,** *1573*. 363-368.

266. Zhou, Q., Shankara, S., Roy, A., Qiu, H., Estes, S., McVie-Wylie, A., Culm-Merdek, K., Park, A., Pan, C., Edmunds, T., *Biotechnol. Bioeng.* **2008,** *99*. 652-665.

267. Guile, G. R., Rudd, P. M., Wing, D. R., Prime, S. B., Dwek, R. A., *Anal. Biochem.* **1996,** *240*. 210-226.

268. Grey, A. A., Narasimhan, S., Brisson, J. R., Schachter, H., Carver, J. P., *Can. J. Biochem.* **1982,** *60*. 1123-1131.

269. Takahashi, N., *J. Chromatogr., A* **1996,** *720*. 217-225.

270. Yagi, H., Takahashi, N., Yamaguchi, Y., Kimura, N., Uchimura, K., Kannagi, R., Kato, K., *Glycobiology* **2005,** *15*. 1051-1060.

271. Arfi, A., Richard, M., Gandolphe, C., Scherman, D., *J. Inherited Metab. Dis.* **2010,** *33*. 61-67.

272. Castaneda, J. A., Lim, M. J., Cooper, J. D., Pearce, D. A., *Acta Neuropathol.* **2008,** *115*. 159-174.

273. Sands, M. S., Vogler, C. A., Ohlemiller, K. K., Roberts, M. S., Grubb, J. H., Levy, B., Sly, W. S., *J. Biol. Chem.* **2001,** *276*. 43160-43165.

274. Takashiba, M., Chiba, Y., Arai, E., Jigami, Y., *Anal. Biochem.* **2004,** *332*. 196-198.

275. Zhou, Q., Kyazike, J., Edmunds, T., Higgins, E., *Anal. Biochem.* **2002,** *306*. 163-170.

276. Leavell, M. D., Kruppa, G. H., Leary, J. A., *Anal. Chem.* **2002,** *74*. 2608-2611.

277. Parolis, L. A. S., Parolis, H., Kenne, L., Meldal, M., Bock, K., *Carbohydr. Res.* **1998,** *309*. 77-87.

278. Takashiba, M., Chiba, Y., Jigami, Y., *Anal. Chem.* **2006,** *78*. 5208-5213.

279. Wang, X.-H., Nakayama, K.-i., Shimma, Y.-i., Tanaka, A., Jigami, Y., *J. Biol. Chem.* **1997,** *272*. 18117-18124.

280. Harvey, D. J., Bousfield, G. R., *Rapid Commun. Mass Spectrom.* **2005,** *19*. 287-288.

281. Du, H., Levine, M., Ganesa, C., Witte, D. P., Cole, E. S., Grabowski, G. A., *Am. J. Hum. Genet.* **2005,** *77*. 1061-1074.

282. Lee, K., Jin, X., Zhang, K., Copertino, L., Andrews, L., Baker-Malcolm, J., Geagan, L., Qiu, H., Seiger, K., Barngrover, D., McPherson, J. M., Edmunds, T., *Glycobiology* **2003,** *13*. 305-313.

283. Dunn, J. D., Reid, G. E., Bruening, M. L., *Mass Spectrom. Rev.* **2009,** *29*. 29-54.

284. Pinkse, M. W. H., Heck, A. J. R., *Drug Discovery Today: Technol.* **2006,** *3*. 331-337.

285. Yu, Y.-Q., Fournier, J., Gilar, M., Gebler, J. C., *J. Sep. Sci.* **2009,** *32*. 1189-1199.

286. Leitner, A., *TrAC, Trends Anal. Chem.* **2010,** *29*. 177-185.

287. Goldberg, D. E., Kornfeld, S., *J. Biol. Chem.* **1981,** *256*. 13060-13067.

288. Varki, A., Kornfeld, S., *J. Biol. Chem.* **1980,** *255*. 10847-10858.

289. Shipley, J. M., Grubb, J. H., Sly, W. S., *J. Biol. Chem.* **1993,** *268*. 12193-12198.

290. Jain, S., Drendel, W. B., Chen, Z.-w., Mathews, F. S., Sly, W. S., Grubb, J. H., *Nat. Struct. Biol.* **1996,** *3*. 375-381.

291. Lau, K. S., Partridge, E. A., Grigorian, A., Silvescu, C. I., Reinhold, V. N., Demetriou, M., Dennis, J. W., *Cell* **2007,** *129*. 123-134.

292. Sleat, D. E., Zheng, H., Qian, M., Lobel, P., *Mol. Cell. Proteomics* **2006,** *5*. 686-701.

# 7.2 Publication list Kenézy Life Sciences Library

DEENK www.lib.unideb.hu
UNIVERSITY AND NATIONAL LIBRARY UNIVERSITY OF DEBRECEN
KENÉZY LIFE SCIENCES LIBRARY

Register Number: DEENKÉTK/148/2013.
Item Number:
Subject: Ph.D. List of Publications

Candidate: Stefan Mittermayr

Neptun ID: N53BUM

Doctoral School: Doctoral School of Molecular Medicine

**List of publications related to the dissertation**

1. **Mittermayr, S.**, Bones, J., Guttman, A.: Unraveling the Glyco-Puzzle:Glycan Structure Identification by Capillary Electrophoresis.
   *Anal. Chem. Epub ahead of Print (2013)*
   DOI: http://dx.doi.org/10.1021/ac4006099
   IF:5.856 (2011)

2. Váradi, C., **Mittermayr, S.**, Szekrényes, Á., Kádas, J., Takács, L., Kurucz, I., Guttman, A.: Analysis of Haptoglobin N-glycome Alterations in Inflammatory and Malignant Lung Diseases by Capillary Electrophoresis.
   *Electrophoresis. Epub ahead of Print (2013)*
   DOI: http://dx.doi.org/10.1002/elps.201300041
   IF:3.303 (2011)

3. **Mittermayr, S.**, Guttman, A.: Influence of molecular configuration and conformation on the electromigration of oligosaccharides in narrow bore capillaries.
   *Electrophoresis. 33* (6), 1000-1007, 2012.
   DOI: http://dx.doi.org/10.1002/elps.201100681
   IF:3.303 (2011)

4. Bones, J., **Mittermayr, S.**, McLoughlin, N., Hilliard, M., Wynne, K., Johnson, G.R., Grubb, J.H., Sly, W.S., Rudd, P.M.: Identification of N-Glycans Displaying Mannose-6-Phosphate and their Site of Attachment on Therapeutic Enzymes for Lysosomal Storage Disorder Treatment.
   *Anal. Chem. 83* (13), 5344-5352, 2011.
   DOI: http://dx.doi.org/10.1021/ac2007784
   IF:5.856



H-4032 Debrecen, Egyetem tér 1. e-mail: publikaciok@lib.unideb.hu

5. **Mittermayr, S.**, Bones, J., Doherty, M., Guttman, A., Rudd, P.M.: Multiplexed Analytical Glycomics:Rapid and Confident IgG N-Glycan Structural Elucidation.
*J. Proteome Res. 10* (8), 3820-3829, 2011.
DOI: http://dx.doi.org/10.1021/pr200371s
IF:5.113

6. Bones, J., **Mittermayr, S.**, O'Donoghue, N., Guttman, A., Rudd, P.M.: Ultra Performance Liquid Chromatographic Profiling of Serum N-Glycans for Fast and Efficient Identification of Cancer Associated Alterations in Glycosylation.
*Anal. Chem. 82* (24), 10208-10215, 2010.
DOI: http://dx.doi.org/10.1021/ac102860w
IF:5.874

**Total IF: 29.305**

**Total IF (publications related to the dissertation): 29.305**

The Candidate's publication data submitted to the Publication Database of the University of Debrecen have been validated by Kenezy Life Sciences Library on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

19 April, 2013

# CHAPTER 8

# Afterword

## 8.1 Keywords

Capillary electrophoresis laser induced fluorescence, ultra performance liquid chromatography, hydrophilic interaction liquid chromatography, oligosaccharides, glycan analysis, structural identification, glycomics, proteomics, biomarker discovery

## 8.2 Abbreviations

| | |
|---|---|
| 2-AA | 2-aminobenzoic acid |
| 2-AMAC | 2-aminoacridone |
| 2-DE | Two dimensional electrophoresis |
| ACP 5 | Acid phosphatase 5 |
| ADCC | Antibody dependant cell cytotoxicity |
| AEC | Anion-exchange chromatography |
| APTS | 8-aminopyrene-1,3,6-trisulphonic acid |
| BGE | Background electrolyte |
| BPE | Base peak electropherogram |
| CDC | Complement dependent cytotoxicity |
| CE | Capillary electrophoresis |
| CHO | Chinese hamster ovary |
| CID | Collision induced dissociation |
| CZE | Capillary zone electrophoresis |
| DP | Degree of polymerization |
| E | Electric field strength |
| EIC | Extracted ion chromatogram |
| Endo H | Endo-ß-N-acetylglucosaminidase H |
| EOF | Electroosmotic flow |
| ER | Endoplasmatic reticulum |
| ERT | Enzyme replacement therapies |
| ESI | Electrospray ionization |
| F | Drag force |
| Fuc | Fucose |
| FucT | Fucosyltransferase |
| G2 | Maltose |
| G7 | Maltoheptaose |

| | |
|---|---|
| GA | Golgi apparatus |
| Gal | Galactose |
| GalT | Galactosyltransferase |
| Glc | Glucose |
| GlcNAc | *N*-Acetylglucosamine |
| GlcNAc-P-T | *N*-acetylglucosamine-phosphotransferase |
| GNT | *N*-acetylglucosaminyl-transferase |
| GU | Glucose units |
| HILIC | Hydrophilic interaction liquid chromatography |
| HPAEC | High pH high performance anion-exchange chromatography |
| HPLC | High performance liquid chromatography |
| IgG | Immunoglobulin G |
| LAL | Lysosomal acid lipase |
| LC | Liquid chromatography |
| LIF | Laser induced fluorescence |
| M | Molecular mass |
| M6P | Mannose-6-phosphate |
| MALDI | Matrix assisted laser desorption ionization |
| Man | Mannose |
| MAN3 | Pauci-mannose |
| MBL | Mannose binding lectin |
| $\mu_e$ | Electrophoretic mobility |
| $\mu_{EOF}$ | Electroosmotic mobility |
| MOAP | Mixed oxide affinity purification |
| MPR | Mannose-6-phosphate receptor |
| MS | Mass spectrometry |
| MS/MS | Tandem mass spectrometry |
| $\eta$ | Newtonian viscosity |
| NeuAc | N-acetylneuraminic acid |
| OST | Oligosaccharide transferase |
| PAD | Pulsed amperometric detection |
| PGC | Porous graphitized carbon chromatography |
| PNGase F | Peptide-N4-(N-acetyl-ß-glucosaminyl)asparagine amidase |
| q | Charge |
| RMT | Relative migration time |
| RP | Reverse-phase chromatography |
| SEC | Size-exclusion chromatography |
| $SLe^X$ | Sialyl Lewis X |
| ST | Sialyltransferase |
| SU | Sugar unit |
| TFA | Trifluoroacetic acid |
| TOC | Total organic carbon |
| ToF-MS | Time of flight mass spectrometry |
| UPLC | Ultra performance liquid chromatography |
| v | Migration velocity |

## 8.3 Acknowledgement

This PhD thesis stands at the end of a challenging road and I would like to express my gratitude to those who accompanied me along this way or kept me on track:

First of all, I would like to thank my family and friends for their encouragement, trust and patience. Especially, I want to thank my parents for their moral support and, above all, for giving me the opportunity to study.

I would also like to gratefully acknowledge all my coworkers at the Horváth Laboratory of Bioseparation Sciences; my supervisor Prof. András Guttman for his guidance, patience and for making it possible to work abroad; Csaba Váradi, Ákos Szekrényes and András Kovács for the teamwork, fun and every single relaxing coffee break.

My deepest gratitude is also due to Prof. Pauline Rudd and all my colleagues at The National Institute for Bioprocessing Research and Training for a great time in Dublin, their priceless help, suggestions, patients and their friendship. I would like to devote special thanks to Dr. Jonathan Bones for being an incredible teacher and a great friend. One thing that's left for him to tell me is the secret source of his seemingly inexhaustible patience. Thanks a million! Furthermore, I would like to thank Dr. T Tharmalingam for her scientific advice, help and for being the probably best bench mate one could have.

Moreover, I want to thank Prof. Barry Karger for his kind assistance and the support of the Barnett Institute. Thank you for giving me the opportunity to join your great team as a visiting scholar at Northeastern University in Boston.

Furthermore, I would like to express my gratitude to Beckman Coulter, PhyNexus and ProZyme for their kind support and providing me with consumables.

The financial support of the University of Debrecen, the K-81839 OTKA grant and the TÁMOP-4.2.2/B-10/1-2010-0024 project, co-financed by the European Union and the European Social Fund, is also gratefully acknowledged.