TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

Electrophoretic analysis of glycoproteins and their interactions with lectins studied by means of microchip capillary electrophoresis in the liquid phase and of gas-phase electrophoretic molecular mobility analyzer at atmospheric pressure

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

unter der Leitung von

Univ.Prof. Mag.pharm. Dr.rer.nat Günter **Allmaier** und

Assoc. Prof. Mag.rer.nat. Dr.rer.nat. Martina **Marchetti-Deschmann**

E 164 - Institut für Chemische Technologien und Analytik

eingereicht an der **Technischen Universität Wien**

Fakultät für Technische Chemie

von

Nicole Yvonne **Engel**, MSc.

1129229

Strohberggasse 31/8, A-1120 Wien

Wien, 28. März 2017

Acknowledgements

Zuerst möchte ich mich bei **Günter Allmaier** für die Möglichkeit bedanken, dass ich meine Arbeit unter seiner Betreuung und in seiner Arbeitsgruppe durchführen durfte. Außerdem danke ich ihm für seine stetige Unterstützung, Förderung und Hilfestellung.
Zusätzlich bedanke ich mich bei **Martina Marchetti-Deschmann** für ihre Betreuung, die wissenschaftlichen Ratschläge, die mich in die richtige Richtung gelenkt haben, und dass sie immer an mich geglaubt und jederzeit ein offenes Ohr für mich hatte.
Daneben danke ich **Victor Weiss** für seine wissenschaftliche Unterstützung, seine Hilfe in jeglicher Situation und die viele Zeit, die er sich immer für mich genommen hat.
Besonders möchte ich mich bei allen dreien für das Verständnis und den Zuspruch auch in schwierigen Phasen bedanken.

Weiterhin bedanke ich mich bei **Susanne Glück**, **Martin Kratzmeier**, **Andreas Rüfer** und **Christian Wenz** von Agilent Technologies für die Zusammenarbeit in diesem Projekt.

Allen derzeitigen und ehemaligen **Mitgliedern der Arbeitsgruppe** danke ich für das ständig freundliche Arbeitsklima, die vielen lustigen Momente und die gute wissenschaftliche Zusammenarbeit. Danke für die schöne Zeit! In einigen von euch habe ich gute Freunde gefunden und ich hoffe, dass wir auch weiterhin noch so viel Spaß haben werden.

Am meisten möchte ich mich von ganzem Herzen bei **Ben**, meinen **Eltern** und meinen **Freunden** für ihre andauernde Geduld, ihr Verständnis und ihren ständigen ermunternden Zuspruch bedanken. Danke, dass ihr mir immer Kraft gegeben habt und immer für mich da wart und seid!

Ohne euch allen wäre all das niemals möglich gewesen.

Danke

Table of contents

Abstract

Glycosylations are considered the most abundant, but also the most complex form of co- and post-translational modifications. Their huge structural and functional diversity renders their analysis rather challenging, as addressed by the present thesis with the validation of two new analytical methods focussing on the electrophoretic separation of glycoproteins in liquid- and gas-phase. Therefore, intact glycoproteins as well as their interactions with lectins were qualitatively and quantitatively investigated by microchip capillary gel electrophoresis (MCGE) and on a nano electrospray gas-phase electrophoretic mobility molecular analyzer (nES GEMMA) instrument.

For glycoprotein analysis MCGE demonstrated high sensitivities, high accuracy in quantitation, and a high reproducibility in sizing, i.e. molecular weight (MW) determination. This makes it a good rapid and easy-to-use alternative to the more elaborate sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Deviations in sizing could be observed with both techniques, SDS-PAGE and MCGE, in comparison to mass spectrometric-derived values. Especially in the case of MCGE, they increased in accordance to the degree of glycosylation. Nevertheless, the combination of MCGE with magnetic bead-based lectin affinity enrichment enabled the targeted analysis of glycoproteins from complex biological samples with high selectivity. Validation of the developed strategy with human serum and mycelia extract of the fungus *Trichoderma atroviride* demonstrated the selective enrichment of glycoproteins. However, the existence of an unspecific binding between the sample and the beads themselves was revealed, too. Affecting glycoproteins, these unspecific interactions can challenge any lectin-based specificity experiment. SDS-PAGE analysis followed by a proteomic approach revealed these results and enabled in addition the identification of two putative glycoproteins for the only fairly studied glycoproteome of *Trichoderma atroviride*.

Based on a size / MW correlation nES GEMMA, on the other hand, allowed MW determinations of the employed glycoproteins and lectins in very good agreement to mass spectrometric values. Operating at ambient pressure and with non-denaturing electrolyte solutions, the system even enabled the analysis of the weak lectin-glycoprotein interactions while maintaining the biological structure of the complex. Interaction studies were performed in regard to binding specificities and affinities towards selected glycoproteins. Results were additionally compared to capillary electrophoresis-on-a-chip-derived data. Next to mere detection, the non-covalent biospecific complex after size-separation could also be sampled onto nitrocellulose membrane with the nES GEMMA device. Subsequent identification by an immunological assay further proved the intact native structure of the analytes throughout the nES process, the gas-phase separation, and even the electrostatic sampling. Consequently, nES GEMMA is a promising platform for the analysis of the weak glycoprotein-lectin interactions specifically with a straightforward sample preparation, a label-free and chemical-

nature independent detection, and a simplified data interpretation as only singly charged species are regarded.

Zusammenfassung

Glykosilierungen zählen zu den häufigsten, gleichzeitig aber auch zu den komplexesten co- und posttranlationalen Modifikationen. Ihre hohe strukturelle und funktionale Vielfalt stellt eine große Herausforderung für viele Analysemethoden dar. Im Rahmen der vorliegenden Arbeit wird versucht, diese schwierige Situation mittels elektrophoretischen Trennungen in flüssiger und gasförmiger Phase zu verbessern. Dafür wurden sowohl intakte Glykoproteine, als auch deren Wechselwirkungen mit Lektinen qualitativ und quantitativ mittels Kapillargelelektrophorese am Mikrochip (MCGE) und mit einem Nanoelektrospray (NanoES) basierenden Ionenmobilitätsspektrometer (gas-phase electrophoretic mobility molecular analyzer, GEMMA) untersucht.

MCGE wies bei der Analyse von Glykoproteinen hohe Sensitivitäten, Genauigkeiten in der Quantifizierung und Reproduzierbarkeiten für die Bestimmung von Molekulargewichten auf. Damit stellt MCGE eine schnelle und leichter zu handhabende Alternative zur aufwendigeren Natriumdodecylsulfat-Polyacrylamidgelelektrophorese (SDS-PAGE) dar. Für beide Methoden, SDS-PAGE und MCGE, konnten allerdings im Vergleich zur Massenspektrometrie Abweichungen hinsichtlich der ermittelten Molekulargewichte festgestellt werden. Diese Abweichungen nahmen besonders im Falle der MCGE mit steigendem Glykosilierungsanteil zu. Dennoch ermöglichte die Kombination von MCGE und Affinitätsanreicherung mit Lektinen, die an magnetische Beads gekoppelten waren, eine gezielte und hoch selektive Untersuchung von Glykoproteinen aus komplexen biologischen Proben. Die hohe Selektivität zeigte sich auch bei der Anwendung der entwickelten Strategie auf Proben von humanem Serum und Myzeliumextrakten des Pilzes *Trichoderma atroviride*. Jedoch wurden zusätzlich unspezifische Wechselwirkungen der Proben mit den magnetischen Beads selbst gefunden, die besonders Spezifitätsuntersuchungen der Probe mit den unterschiedlichen Lektinen beeinflussten. Die erhaltenen Ergebnisse wurde letztendlich mittels SDS-PAGE und Proteomstudie bekräftigt, wobei zusätzlich zwei Glykoproteine für das bis dahin kaum erforschte Glykoproteom von *Trichoderma atroviride* identifiziert werden konnten.

Im Gegensatz dazu ermöglichte nES GEMMA basierend auf einer Größe-Molekulargewichts-Korrelation Molekulargewichtsbestimmungen der Glykoproteine und Lektine, die sehr gut mit den massenspektrometrischen Werten übereinstimmten. Da das System unter Atmosphärendruck und mit nicht-reduzierenden Elektrolytlösungen arbeitet, konnten sogar die im Vergleich zu anderen Biomolekülen nur schwach wechselwirkenden Lektin-Glykoprotein-Komplexe in ihrer nativen biologischen Form untersucht werden. Interaktionsstudien erlaubten Rückschlüsse auf die Bindungsspezifitäten und -affinitäten der verwendeten Lektine gegenüber den untersuchten Glykoproteinen. Die Interaktionen wurden zusätzlich kapillarelektrophoretisch am Mikrochip untersucht. Neben der bloßen Detektion konnten die nichtkovalenten biospezifischen Komplexe mit

nES GEMMA nach der elektrophoretischen Separation auf eine Nitrozellulosemembran gesammelt werden. Ein anschließender immunologischer Assay belegte die Aufrechterhaltung der nativen Struktur des Biokomplexes während des nES-Prozesses, der Auftrennung in der Gasphase und sogar während des elektrostatischen Probensammelns. Das nES GEMMA-System liefert daher einen vielversprechenden Ansatz für die Analyse von schwachen Lektin-Glykoprotein-Interaktionen. Dabei zeichnet es sich vor allem durch eine einfache Probenpräparation, eine markierungsfreie Detektion, die unabhängig von der chemischen Natur des Analyten ist, und eine vereinfachte Dateninterpretation aufgrund der Berücksichtigung von nur einfach geladenen Spezies aus.

Introduction

1. Glycoproteins

It is estimated that the number of co- or post-translationally modified proteins in eukaryotes significantly exceeds unmodified ones, with the most abundant modification being glycosylations [1]. Glycosylation is regarded the most complex modification type due to the huge heterogeneity of glycoforms created during the expression of glycoproteins as well as the high diversity of functional purposes. Glycoproteins occur in almost all forms of life and can be found inside cells, in extracellular fluids, as well as in cell membranes. They are carrying carbohydrate chains in various degrees, all covalently linked to the polypeptide backbone of the protein. The attached glycan moieties can differ from single monosaccharides to complex linear or branched oligo- or even large polysaccharides. Additional diversity is created by microheterogeneity, which describes the preferred attachment of different glycan moieties to a certain glycosylation site of a protein. This leads to the formation of the already mentioned glycoforms, which can be further modified at altering degrees (e.g. by sialylation, acetylation, acylation, sulfation, fucosylation or methylation). In this regard, the generation of such glycoforms is rather regulated and can be dependent on the cell state, disease or function of an organism.

1.1 Occurrence and structure of glycans

Most common carbohydrates are attached N-glycosidically or O-glycosidically between the reducing end of the glycan chain and the amide group of asparagine (N-glycans) or the hydroxy group of serine, threonine, hydroxyproline, hydroxylysine, and tyrosine (O-glycans). Additionally, linkages via ethanolamine phosphates (glycosyl-phosphatidylinositol anchor), phosphodiester bridges, and C-glycosides can be found [2].

1.1.1 N-glycans

The synthesis of N-glycans is a combination of co- and post-translational processes and, despite the great diversity of glycosylations, the result of a highly conserved pathway, as described in the following paragraphs. During the eukaryotic biosynthesis of the polypeptide chain, a pre-synthesized triantennary precursor consisting of three glucoses, nine mannoses, and two N-acetylglucosamines ($Glc_3Man_9GlcNAc_2$) is transferred *en bloc* onto the growing protein in the lumen of the endoplasmic reticulum (ER). This tetradecasaccharide is then modified in the ER and Golgi apparatus in subsequent trimming and attachment processes catalysed by several glycosidases and glycosyltransferases.

For the precursor formation monosaccharides are successively added by glycosyltransferases onto the lipid carrier dolichol-pyrophosphate, which is located in the ER membrane [3]. The enzyme oligosaccharyltransferase (OT) then transfers the glycan onto the polypeptide [4, 5], where the N-acetylglucosaminyl(β1–N)asparagine bond is formed [6]. For that the OT has to recognize a certain glycosylation sequon on the still unfolded polypeptide, which is formed by asparagine, any second natural amino acid except Pro, and Ser or Thr (Asn-X-Ser/Thr). Furthermore, this motif is required to be able to adopt a loop conformation (Asx-turn) in order to be a substrate for glycosylation by the OT [7]. This so-called Asx-turn causes a nucleophilic enhancement of the asparagine amide nitrogen by bringing Ser/Thr in closer proximity, as proposed in two different mechanistic models by Imperiali et al. [8, 9] and Bause et al. [10, 11] (**Figure 1.1**). Yet, not all available sequons of a protein are glycosylated, depending on numerous factors like the amino acids within and adjacent to the sequon, the sequon location in the whole polypeptide chain, the protein folding state, as well as the composition of the OT subunits and the availability of dolichol carriers [12].

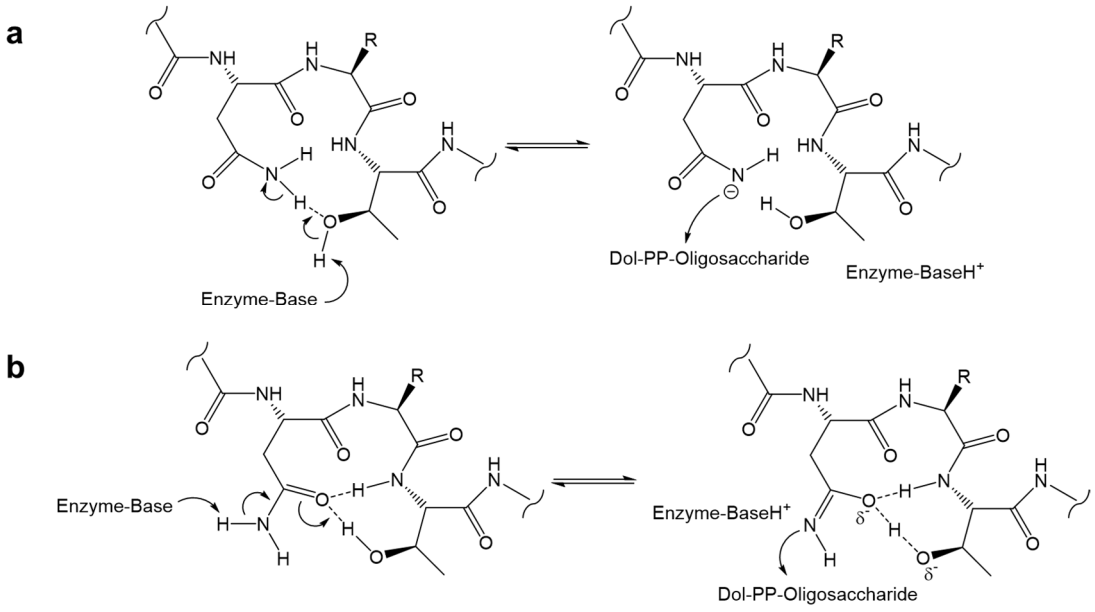


Figure 1.1 – Mechanistic models of asparagine activation for OTs. Mechanisms involve either amide deprotonation ([11], **a**) or amide tautomerizaion ([9], **b**).

N-glycosylation in Gram-negative bacteria and archaea follows a similar process, only Gram-positive bacteria seem to show no signs of N-glycans [13]. In both former cases *en bloc* transfer of an oligosaccharide precursor by OTs from lipid carriers (dolichol in archaea and undecaprenol in bacteria) are observed. However, the precursor is not characterized by a conserved structure as in the case of the eukaryotic tetradecasaccharide. Moreover, in eukaryotes proteins are glycosylated before their folding, whereas they are presumed to be fully folded in prokaryotes. Although glycosylation requires

a similar consensus sequence in all species, the actual rate of glycosylation for bacteria is believed to be far more restricted [14].

Next to that, an additional pathway was discovered in the bacterium *Haemophilus influenza* performing cytoplasmic N-glycosylation [15]. Here, glucose and galactose are independently transferred to the asparagine of the N-glycosylation consensus sequence by a N-linking glycosyltransferase forming mono- and dihexoses.

In eukaryotes the transfer of the oligosaccharide precursor is followed by a post-translational trimming process as well as further modifications by glycosylhydrolases and glycosyltransferases in the ER and the Golgi apparatus leading to the huge structural variety of glycan moieties. However, they all share a common core consisting of three mannoses attached to a chitobiose core: Man(α1–6)[Man(α1–3)]Man(β1–4)GlcNAc(β1–4)GlcNAc. To this pentasaccharide further monosaccharides can be added, which can result in whole carbohydrate chains (antennae, **Figure 1.2**). Based on the degree and composition of branching as well as the variety of monosaccharides attached, three main groups of N-glycans can be differentiated: high-mannose, complex, and hybrid type (**Figure 1.3**).

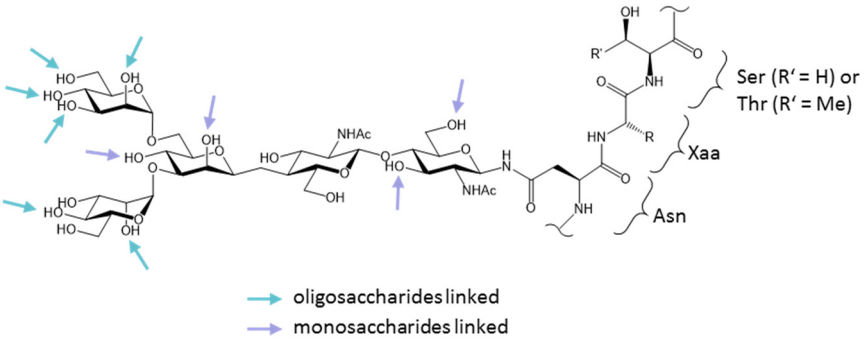


Figure 1.2 – Core pentasaccharide in N-glycans. The conserved trimannosyl-chitobiose core is linked to the consensus sequence Asn-Xaa-Ser/Thr and can be further modified by attachment of oligosaccharide chains or monosaccharides, as indicated by turquoise and purple arrows for possible points of attachments (according to [16])

The high-mannose type is characterized by usually two to six mannose residues bound to the core, but even higher numbers can be found e.g. in the case of bovine thyroglobulin (nine mannose residues, [17]) or in proteins produced by yeast (100 to 200 mannose residues, [18]). Complex type glycans can carry up to five antennae of one or several N-acetyllactosamine units (Gal(β1–4)GlcNAc), which are often end-capped by sialic acids. Also fucosylation and additional monosaccharides in the core or antennae can be found. The hybrid type is a mixture of both other types having mannose residues as well as complex type antennae attached. Several factors, which can influence the type of glycosylation,

are for instance the amino acids adjacent to the glycosylation site, the accessibility of the glycans to the processing enzymes, and the cell type, in which the protein is produced [19].

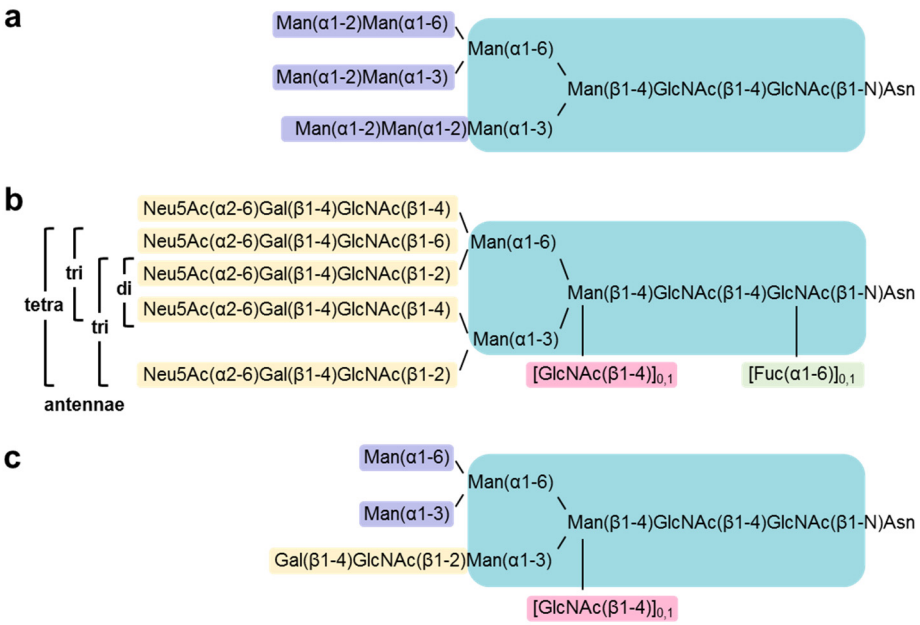


Figure 1.3 – Main N-glycan types. Depending on the type and degree of mono- or oligosaccharides attached to the core structure (turquoise), several different types of N-glycans can be classified: high-mannose (**a**), complex (**b**), and hybrid type (**c**). N-acetyllactosamines capped with sialic acids are represented in yellow, mannose residues in purple, N-acetylglucosamines in red, and fucose in green. (according to [16])

1.1.2 O-glycans

In comparison to N-glycans, O-glycosylation in eukaryotes is mainly a post-translational process in the Golgi apparatus with the help of transferases concerning the final and folded protein. O-glycans are created by the successive attachment of monosaccharides, each offering three to four possible linkage sites in α- or β-configuration leading to different branches and compositions [20]. Moreover, they vary in their protein-glycan linkages, which further results in extremely diverse structures (**Figure 1.4**). O-glycans are classified according to the first monosaccharide that is linked to the hydroxy amino acid of the protein backbone (e.g. Ser, Thr, Tyr, hydroxylysine Hyl, or hydroxyproline Hyp).

A very common form of O-glycans in nature are the so-called mucin-type glycoproteins, in which a N-acetylgalactosamine is α1–O linked to Ser or Thr of a protein. Eight different mucin-type core structures can be distinguished in accordance to the second bound sugar and its linkage (**Table 1**) [21]. Subsequently those core structures can be specifically elongated or modified by e.g. sialylation, sulfation, acetylation, fucosylation, and poly-N-acetyllactosamine-extension leading to a large number of structures [2]. Usually they consist of one to 20 monosaccharide residues.

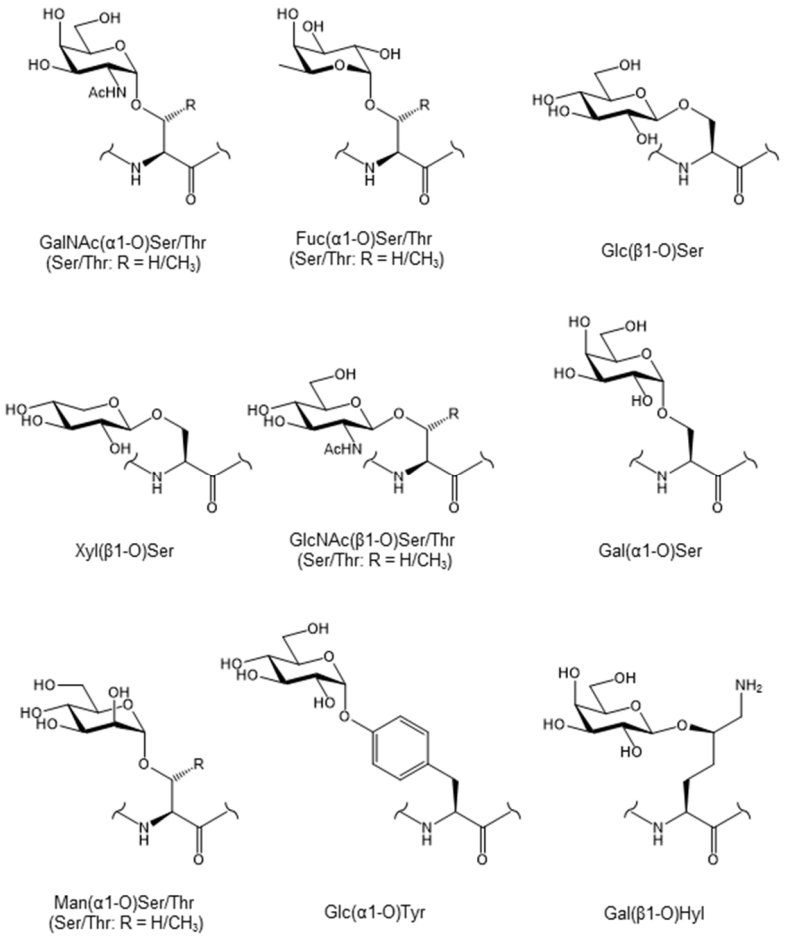


Figure 1.4 – Exemplary O-glycan structures. O-glycans can be classified according to the first monosaccharide (e.g. GalNAc, Fuc, Glc, Xyl, GlcNAc, Gal, Man) and the amino acid (e.g. Ser, Thr, Tyr, Hyl, Hyp), to which the monosaccharide is α- or β-O-glycosidically linked to.

mucin core type	structure
1	Gal(β1-3)GalNAc
2	Gal(β1-3)[Gal(β1-6)]GalNAc
3	GalNAc(β1-3)GalNAc
4	GalNAc(β1-3)[GalNAc(β1-6)]GalNAc
5	GalNAc(α1-3)GalNAc
6	GalNAc(β1-6)GalNAc
7	GalNAc(α1-6)GalNAc
8	Gal(α1-3)GalNAc

Table 1 Core region sequences of O-glycans as found in serum, cell membrane, and mucine-type glycoproteins. (according to [21])

In contrast to N-glycans no general consensus sequence for mucine-like O-glycosylations could be determined so far. However, certain predictions can be made [2]. Thus, glycosylation is dependent on the specificity and activity of the tissue-specifically expressed glycosyltransferases. Moreover, as only

folded proteins are targeted, merely Ser/Thr exposed on the protein surface can be glycosylated. They are usually in regions of extended conformations (e.g. β-turns, rich of Pro), low hydrophobicity, and low steric hindrance (no large amino acid). Moreover, O-glycosylated Thr and Ser often have Ser, Thr, Pro, and, in the case of Thr, Val residues in close proximity, whereas Met, Asp, and Asn are rarely found. O-glycosylations, in contrast to N-glycans, often occur in close proximity on short peptide sections comprising repeating units of Ser, Thr, and Pro.

Next to mucine-type O-glycosylations, O-linked fucose (Fuc) and glucose (Glc) in epidermal growth factor domains [22-24], GlcNAc in cytosolic and nuclear proteins [25], and, to a lesser amount, mannose in mammalian brain, nerves, and skeletal muscles have been described [26]. Similarities of O-linked GlcNAc with phosphorylations have been suggested, as both processes are reversible and dynamic and show a rapid response to extracellular stimuli [27]. Thus, these types of O-glycosylations are assumed to also take part in regulatory functions. Galactose (Gal) linked to Ser has been found in plant glycoproteins, on collagen domains, and in glycoproteins from cellulosome, which is an extracellular complex of cellulases in cellulolytic bacteria [28]. O-linked xylose (Xyl) has only been identified in proteoglycans so far [29].

Yeast, which mainly creates oligomannose structures, can be considered an exception in eukaryotic O-glycosylation. In contrast to mammalian cells the biosynthesis is located in the ER, which enables the possibility of co-translational O-glycosylation in the unfolded or partially folded protein. Moreover, the sugars exhibit a different structure with dolichol-phosphate-mannose being the sugar donor and the sequence specificity is very different with glycine and proline having inhibitory effects on glycosylation [30].

Some Gram-negative bacteria have shown the ability for OT-dependent O-glycosylation, which resembles very much N-glycosylations mediated by OTs [31, 32]. This pathway involves the *en bloc* transfer of oligosaccharides from a lipid carrier onto the protein, however without any consensus sequence. In Gram-positive bacteria and archaea, on the other hand, a processive O-glycosylation as in eukaryotes is prevalent, yet with a higher O-glycan abundance in bacteria than in archaea [13]. Generally, N-glycans are rather predominant in archaea, whereas O-glycans appear more commonly in bacteria. Compared to eukaryotes, however, they both offer a much greater diversity of glycosylation concerning the range of monosaccharides as well as their types of linkages and modifications.

1.2 Functions of protein glycosylations

The structural diversity of glycosylations is also reflected in the variety of functional purposes glycans can fulfil or be part of (details see e.g. [33, 34]). Those functions are mostly of structural or modulatory

nature and / or they concern molecular recognition events. Furthermore, they span from being trivial to being crucial for growth, development, or survival of an organism.

1.2.1 Structural and modulatory functions

Glycosylations can influence the physical, chemical, and biological properties of a protein by changing mainly its size, structure, conformation, charge, solubility, stability, or its dynamic properties. For instance, the bulky glycans can protect the protein from recognition by proteases and antibodies, which are proteolyzed only after removal of the glycosylation [35]. Moreover, they can increase structural rigidity. Additionally, glycans can modify or fine-tune the activity of enzymes and signalling molecules and play key roles in protein expression and processing.

Most of the time the effects of glycosylations on a protein are not distinguishable, but rather concerted. Thus, changes in the glycoprotein structure or its outer charge can also influence e.g. its biological activity or stability. For example, high degrees of sialylation or sulfation may increase a glycoproteins overall charge and, consequently, increase solubility. Often the same glycan has different functions at distinct locations, at specific times, or after certain external stimuli within a single organism.

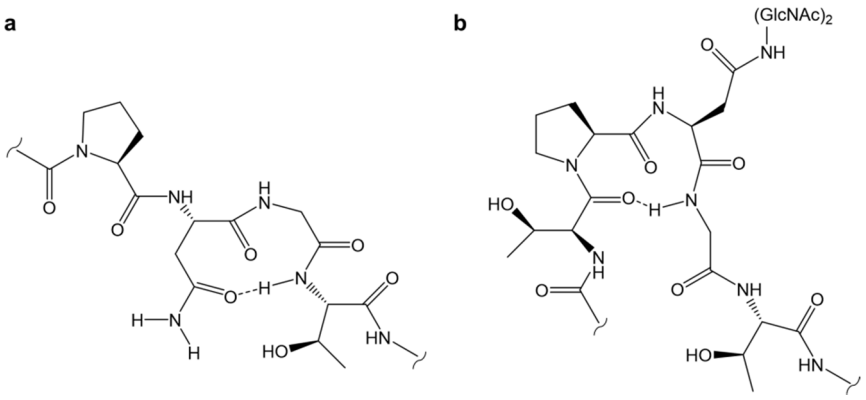


Figure 1.5 – Influence of glycans on the protein folding. N-glycosylations can affect the folding of the polypeptide chain close to the glycosylation site and induce a conformational switch from Asx-turn (**a**) to a type I β-turn (**b**) [36].

N-Glycosylations can have an essential role in the correct folding of the polypeptide chain adjacent to the glycosylation site. It was proposed that they support the transformation from an extended Asx-turn to the more compact β-turn (**Figure 1.5**) [36]. However, this conformation might only be a transient structural change during the folding process and is not always maintained in the final glycoprotein [37]. Other kinetic folding studies suggest that glycosylations decrease the unfolding rate of the native state [38]. To some degree the last finding was supported by folding studies of natively

glycosylated erythropoietin (EPO) compared to a non-glycosylated variant [39]. During the three-state folding mechanism of EPO, the glycans did not alter the folding mechanism itself but slowed its rate and increased the stability of the intermediate species, biasing certain folding pathways.

Like N-glycans, also O-glycans affect the secondary, tertiary, and quaternary structure of the protein [2, 20]: they can interrupt α-helical peptide sections, trigger conformational changes of the protein, or enable the formation of protein aggregates. Eliminating, for instance, all O-glycosylations of ovine submaxillary mucin causes a drastic conformational transformation from linear to globular shape. Moreover, the mucin typical filamentous structure are only formed due to the many O-glycans and the resulting multiple interactions between the Ser/Thr-linked N-acetylgalactosamine and the peptide backbone.

In another instance, glycosylations can be able to protect organisms from heat or cold. Antifreeze glycoproteins in the blood of Antarctic fish decrease the freezing point of the fish by binding to ice crystals and preventing their nucleation and growth inside the fish [40]. Therefore, the fish does not freeze in the icy environment. As the other extreme, some heavily glycosylated glucanases in *Saccharomyces cerevisiae* showed a comparably much higher heat stability than non-glycosylated counterparts expressed in *Escherichia coli* [16].

Furthermore, glycosylations can affect the biological activity and dynamic properties of a protein. The efficiency of for example bovine pancreas ribonuclease (RNase) is directly influenced by its degree of glycosylation. With a growing glycan chain at its only glycosylation site Asn-34, also the dynamic stability of RNase and, additionally, the steric hindrance between the glycans and the substrate of the enzyme (double stranded RNA) is increasing. This causes a drastic decrease in activity, especially compared to the non-glycosylated RNase [41].

1.2.2 Involvement in recognition events and interaction with carbohydrate recognizing proteins

Next to physicochemical consequences for glycoproteins, glycans participate in various physiological and pathological recognition events (**Figure 1.6**). They act as specific receptors for microorganisms, toxins, or antibodies, take part in intra- and intercell trafficking, and mediate cell-matrix or cell-cell interactions. Moreover, carbohydrate moieties can help to orientate proteins, e.g. relative to a membrane [42], or to direct a protein to a specific location in the cell, e.g. leukocytes to sites of inflammation [43]. Most of the times recognition events involve carbohydrate-binding proteins (CBPs), which either recognize glycans from a different (extrinsic) or from the same organism (intrinsic) and mainly mediate cell-cell interactions as well as interactions with extracellular molecules [34]. Typical CBPs are found in the large protein groups of antibodies, lectins, or enzymes involved in modifying glycans.

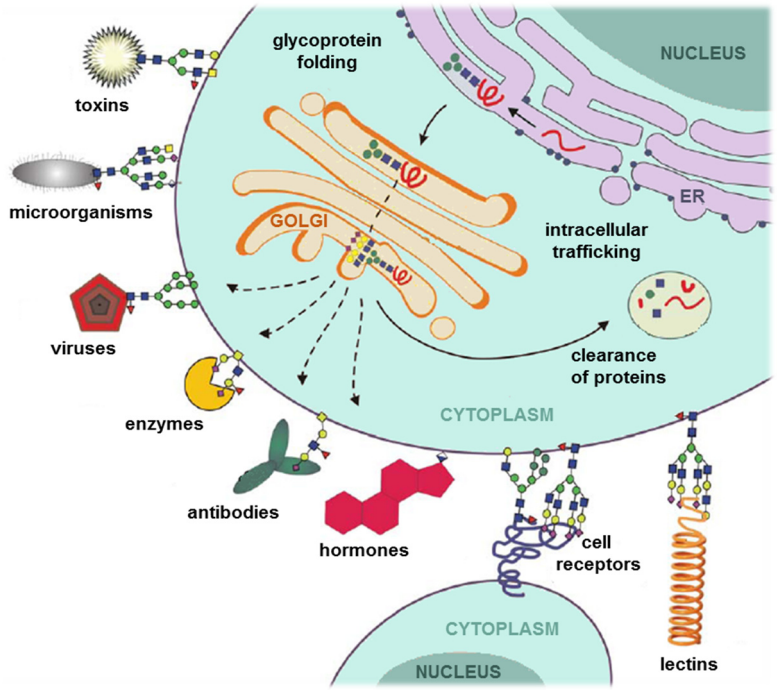


Figure 1.6 – Glycoprotein recognition. Glycoproteins are involved in a variety of binding processes with different interaction partners like viruses, enzymes, antibodies, or lectins, and play a role in many intra- and intercell recognition events (according to [44]).

Enzymes as CBPs

Carbohydrate recognizing enzymes in the form of glycosidases and glycosyltransferases already play a role in the synthesis of glycoproteins. Further enzymes are responsible for additional modifications of glycans like phosphor-, sulfo-, O-methyl- or O-acetyltransferases, among others. Glycosidases remove monosaccharides from a sugar chain as they do with the precursor oligosaccharide in the post-translational trimming process of N-glycans. Glycosyltransferases, on the other hand, sequentially transfer sugar moieties (monosaccharides, disaccharides, or oligosaccharides) from an activated donor substrate (e.g. nucleotide sugars, lipids) onto acceptor substrates (e.g. mono- or oligosaccharides, proteins, lipids, small organic molecules, or DNA) creating linear or branched polysaccharides [45]. The new glycosidic bond can be formed either in a retaining or inverting mechanism concerning the anomeric configuration of the donor substrate. Similar mechanisms were observed for glycosidases (**Figure 1.7**).

Antibodies as CBPs

One very early example for antibodies that can recognize certain carbohydrate structures was already discovered 1900 by Karl Landsteiner working in the Pathological Anatomy Institute of the University of Vienna (Austria) [46]. He revealed that human sera of some individuals agglutinated the red blood cells of others and that the agglutination was not due to bacterial antibodies or antibodies formed in the

response of an infection, as formerly believed [47]. Landsteiner further introduced the three human blood groups stating that the erythrocytes of individuals with blood group A or B carry antigens A and B, respectively, or none in case of the third blood group (later called 0). Moreover, the human serum does not contain the corresponding antibody (agglutinins anti-A and anti-B) for the own antigen present, whereas both can be found in blood group 0. A fourth group AB with both antigens expressed on the red blood cells and without any antibodies in the serum was discovered later. Antigens A and B, as well as the additionally observed antigen H on 0 group erythrocytes, which was proposed to be a precursor for antigens A and B [48], were also found to occur in human exocrine secretions and in mucosal tissues of many animals. But it was only until later that the blood group determinants could be identified as carbohydrates of the mucin-type [49, 50].

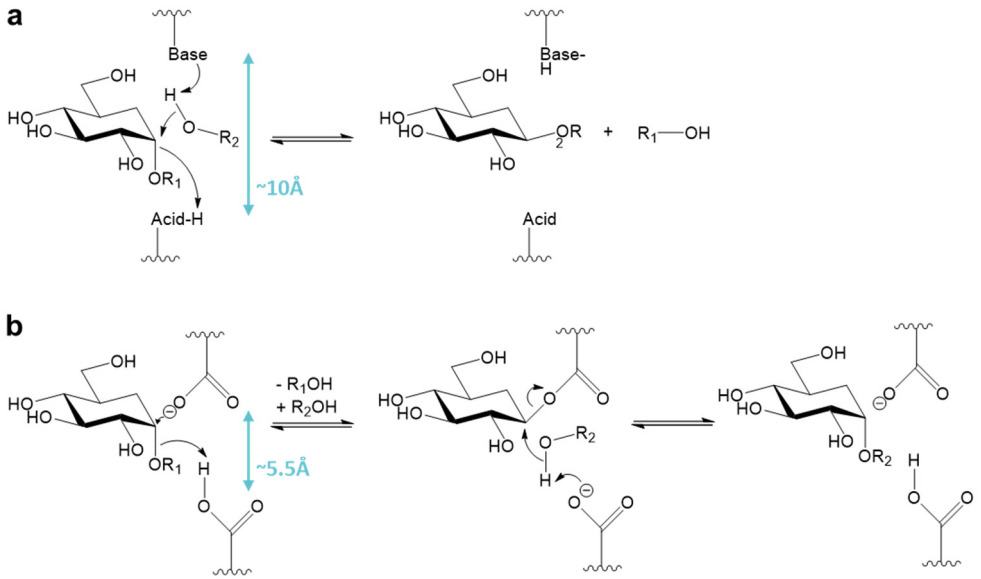
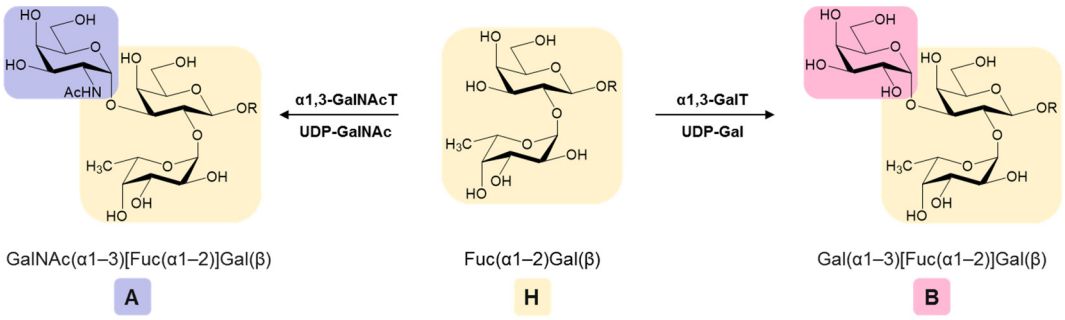


Figure 1.7 – Proposed mechanisms for glycosidases and glycosyltransferases. Reactions result in an inversion (a) or the retention (b) of the anomeric configuration involving S_N2-like attacks and oxocarbenium ion-like transition states [45]. The retention requires a second successive S_N2-like reaction after the formation of a glycosyl-enzyme intermediate. In the case of glycosidases, R_1 corresponds to a proton and R_2 to a glycan remainder. For glycosyltransferases, R_1 is typically a nucleoside mono/diphosphate and belonging to the donor substance and R_2, as part of the acceptor group, e.g. either another sugar or a protein.

Today the synthesis and structure of the antigen determinants is established (**Figure 1.8**). Certain glycosyltransferases build the final antigens onto different precursor saccharides of glycolipids or proteins. Their peripheral disaccharides are the basis for the differentiation into the antigen types. Fucosylation of the precursor by an α1–2 fucosyltransferase (FucT) creates the H antigen determinant, which can be further modified by N-acetylgalactosaminyltransferases (GalNAcT) or galactosyltransferases (GalT) resulting in the antigen determinants for A and B, respectively.



antigen type	peripheral disaccharide core structures
type-1	Gal(β1–3)GlcNAc(β) – (poly-)N-acetyllactosamines
type-2	Gal(β1–4)GlcNAc(β) – (poly-)N-acetyllactosamines
type-3	Gal(β1–3)GalNAc(α) – O-GalNAc glycans
type-4	Gal(β1–3)GalNAc(β) – on glycolipids

Figure 1.8 – Minimal determinant structures of blood group antigens A, B, and H. Antigens A and B can be created by further modification of the antigen H. The antigen types depend on the peripheral disaccharide structures, to which the determinant blood group antigens are linked.

Next to these antibodies directed to blood antigens, several other natural anti-glycan antibodies can be found in human tissues. For instance, anti-Gal antibodies have become a major focus concerning the topic of xenotransplantation (transplantation of organs from different species like pigs to humans), as they are responsible for hyperacute rejections of the organs. They are targeting specific carbohydrate epitopes (xenoantigens) of the pig, e.g. Gal(α1,3)Gal(β1-4)GlcNAc (the Gal-antigen) on the vascular endothelial cells of the transplanted organ [51]. Within minutes the binding of the antibodies to the antigens can cause irreversible graft damage by inducing the complement cascade. This includes congestion and thrombosis in the small vessels of the graft, followed by disruption of the vascular endothelium as well as massive interstitial haemorrhage and oedema – very similar consequences as known from AB0-incompatible allotransplantations (transplantations within same species). The Gal-epitope has been observed in many non-primate mammals, also very abundantly in pig tissues, but not in Old World monkeys, apes, or humans. The latter ones, on the contrary, produce anti-Gal antibodies supposedly in response to microorganisms in their gastrointestinal tract in the first months of life [52]. They were found in immunoglobulin subclasses IgM, IgG, and IgA, of which mainly IgM and IgG are involved in hyperacute rejection processes. Next to the Gal-epitope also other xenoantigens in pigs are supposed to be involved in the rejection process, e.g. N-glycolylneuraminic acid (Neu5Gc(α2-3)Gal(β1-4)Glc), the so-called Hanganutziu-Deicher antigen (HD-antigen). As with the Gal-epitope before, the HD-antigen is widely distributed amongst mammals except in humans, who in turn possess anti-HD antibodies [53].

Lectins as CBPs

Lectins are a very large and diverse group found in all living organisms throughout nature. They have become a versatile tool in the detection, isolation/purification, and characterization of glycoproteins as well as the analysis of cell-surface glycans, whose changes in pattern are often related to diseases and cancer [43]. Moreover, lectins have shown to influence many crucial cellular processes due to their interactions with carbohydrates. They can e.g. mediate bacterial and viral adhesion to host cells, control the intracellular traffic of glycoproteins, induce cell growth, mitosis, or apoptosis, or regulate cell-cell interactions.

First hints that proteins have the ability to agglutinate erythrocytes were already mentioned at the end of the 19th century: Stillmark and Helling each isolated a highly toxic hemagglutinin from the castor tree (ricin) and the jequirity bean (abrin), respectively [43]. The high toxicity of ricin led to several attempts to apply it as biochemical weapon. As known today, both plant lectins are heterodimeric complexes consisting of a N-glycosidase part and a galactose-specific agglutinin part. Further evidence for a protein being able to agglutinate erythrocytes and other cells was gained with the isolation of concanavalin A (ConA) from jack bean (*Canavalia ensiformis*) by Sumner and Howell [54]. They could also demonstrate sugar specificity of the lectin for the first time and proposed a connection between hemagglutination and interactions of ConA with carbohydrates on the erythrocyte surface, resembling antibody (blood agglutinins) interactions as described above. In relation to the identification of the different blood types, also a blood type specificity of various plant agglutinins was detected [55, 56], which finally gave them their general name 'lectin' (lat. *legere*, to choose, select, or pick out) for all sugar-specific agglutinins of nonimmune origin [57].

Although plant lectins were the first to be discovered, only little is known about their natural functions and ligands. In contrast, the first isolation of animal lectins was half a century later (1952 in eel [49]), yet they are studied today in much more detail. Generally, lectins are classified according to homologies in primary structure and relationship in evolution. Thus, for instance at least 13 different lectin families of structurally and evolutionary related proteins could be classified in animals and 7 in plants (**Table 2**) [58-61]. However, these classifications are not always straightforward and universal. There are some structurally unique lectins, which are not part of a greater lectin family, but of protein families not associated with carbohydrate recognition. Next to that, lectin families can include proteins either without the ability to bind sugars or with the ability to (additionally) bind other molecules like proteins, lipids, or nucleic acids [58, 59]. Furthermore, some lectin families are more diverse and, therefore, further divided into sub-groups according to gene structure and nature of additional non-lectin domains as in the case of the C-type (Ca^{2+} requiring) lectins. This group of lectins recognize very different sugar moieties and only share a lectin module.

	structural lectin families	carbohydrate ligand
animal:	**C-type** (including calcium-dependent, including selectins, collectins, endocytic lectins)	various
	galectins (former S-type)	β-galactosides
	P-type (phosphomannosyl receptors)	Man-6-phosphate on high-mannose type N-glycans
	I-type (including immunoglobulin superfamily members, siglec family)	variable (singlecs: sialic acid)
	calnexin, calreticulin, calmegin	glycosylated high-mannose type N-glycans (Glc_1Man_9)
	hyaluronan-binding proteins	hyaluronan chains
	M-type (α-mannosidase-related lectins)	high-mannose (Man_8)
	L-type	various
	R-type	various
	F-box	$GlcNAc_2$
	ficolins (fibrinogen-type)	GlcNAc, GalNAc
	chitinase-like lectins	chito-oligosaccharides
	F-type	Fuc-terminating oligosaccharides
	intelectins	Gal, galactofuranose, pentoses
plant:	**amaranthins**	GalNAc, T-antigen (Gal(β1-3)GalNAc)
	chitin-binding composed of **haevin** domains	chitin, GlcNAc-trimers and -tetramers
	Cucurbitaceae phloem	GlcNAc-oligomers
	jacalin-related	terminal Gal or Man/maltose
	L-type (legume lectins)	various
	monocot mannose-binding (**MMBL**)	Man-oligomers
	R-type (type 2 ribosome-inactivating proteins (type 2 RIP))	various

Table 2 Exemplary animal and plant lectin families with structural or evolutionary homologies. (according to [34, 58, 60, 61])

Despite the great variety of the lectins in size, primary structure, and composition, some similarities in regard to their tertiary and quaternary structure as well as carbohydrate binding sites (CBS) were observed independently from their source [60]. Comparing lectins of different families, usually no common fold is expected and also the CBSs are rather diverse regarding the high numbers of different ligands and biological functions. Within families like galectins, however, which all have a specificity towards the same oligosaccharide, the CBS is highly conserved [62]. In contrast, C-type lectins show many variations within their own family due to diverse specificities [63]. Nevertheless, common features have been revealed comparing lectins from different families and even different origins. One example is the legume lectin fold, a β-sandwich of antiparallel β-strands, which was first discovered in the lectin ConA and is typical for legume lectins [43]. Although a detailed analysis shows certain differences as the number of involved β-strands, a similar fold could be found in animal lectins like galectins, pentraxins, and ERGIC-53 [60]. Other examples of typical lectin folds, which can be both detected in plant and animal lectins, are the β-trefoil fold (three subdomains of each four-stranded antiparallel β-sheets, identified in e.g. ricin-like lectins and fibroblast growth factors [64, 65]) and the

hevein-domain (small disulfide-rich domains e.g. in wheat germ agglutinin WGA and cobra venom cardiotoxin [66, 67]).

Some lectins like those of the C-type and the legume lectins are dependent on divalent metal ions, which influence local conformational changes and thus the functionality of the CBS. Both lectin families have a primary CBS selective for a certain monosaccharide and additional CBSs with fewer interactions to a carbohydrate residue covalently linked to the primary monosaccharide. The whole CBS consists of conserved residues providing hydrogen bonds for interaction with the sugar moiety and two variable loops for additional hydrogen bond and van der Waals interactions [60]. These determine the specificity of the primary binding site and the subsites.

Single-site binding affinities between a lectin and its carbohydrate are rather low (with dissociation constants in the micro- to millimolar range), especially compared to the very strong antibody-antigen interactions (nano- to picomolar range) [68, 69]. Binding is non-covalent and includes typically hydrogen bonds, but also hydrophobic, van der Waals, and electrostatic interactions [70]. In order to increase the avidity, lectins usually make use of multivalence, which describes the simultaneous interaction of the mostly multiple epitopes of the ligand with several binding sites of the lectin. Therefore, the lectins are either multidomain proteins offering more than one CBS per lectin monomer or they form multimeric proteins [71]. Variations in the quaternary structure can also influence the lectin specificities depending on the multivalent ligand and the locations of its glycosylation sites towards each other. Moreover, binding kinetics of multivalent interactions can vary in comparison to monovalent ones, which can be used for fine-tuning certain binding processes [72].

2. Analysis of glycoproteins

Differences in glycosylation pattern can impact development, differentiation, physiology, and disease of cells or organisms and in turn allow to draw conclusions about their state. Defects in the synthesis or attachment of glycans, as known from more than 40 genetic diseases in humans (congenital disorders of glycosylations), can also have severe consequences on the nervous system with no effective treatment known to date [44]. Moreover, changes in the surface glycans of cells have been connected with the development of cancer [43, 73]. So, high sialylation or fucosylation, altered sulfations, high branching degree of N-glycans and their extension with polylactosamines have been found specifically in tumour cells. Often, the identification of such differences can be used for detecting disease biomarkers in clinical studies, the development of treatments or prognosis. Many therapeutic proteins today are complex glycoproteins, typically produced as recombinant glycoproteins in a variety of expression systems (the company first produced and patented such a glycoprotein is termed originator). Often these biotherapeutics, after the patent-protection has ended, are hitting the market as so-called biosimilars [74]. As the glycan chains of those biotherapeutics heavily influence their stability, biological activity, immunogenic potential, safety, and pharmacodynamic properties, glycosylations are often optimized or tried to be made with the same glycan pattern. However, this requires a controlled environment and production has to be monitored closely. Also batch-to-batch variations caused by smallest changes during cultivation (temperature, pH, composition of the medium, addition of nutrients/growth factors/hormones, etc.) have to be monitored. Is the biosimilar a glycoprotein, it usually consists of a variety of different glycoforms as expected. When introduced to the market, no significant additional glycoforms are allowed by the governmental control organization.

This all emphasizes the importance of finding reliable and rapid methods of analysis. Due to the huge structural complexity of glycoproteins and their glycosylations there is still no universal method of choice for their investigation. Usually a combination of different analytical techniques has to be applied for satisfying results depending on the information of interest (e.g. identification of glycosylation sites, structural elucidation of glycans, interaction with other proteins, functional properties).

2.1 Investigation of intact glycoproteins

Generally, there are three main approaches in the analysis of glycosylations: (i) investigation of the intact glycoprotein in its native or denatured conformation, (ii) examination of glycopeptides, and (iii) characterization of chemically or enzymatically released N- and O-glycans [44, 75-78]. The following part, however, will concentrate on exemplary analytical methods concerning the investigation of intact glycoproteins. However, most of those techniques can also be applied for glycopeptide or glycan analysis. Many approaches involve techniques like capillary electrophoresis (CE), liquid

chromatography (LC), mass spectrometry (MS), as well as their combinations and variations. Structural elucidation of an intact glycoprotein can be challenging, as it usually exists as a mixture of glycoforms and often only small amounts of isolated and purified material are available. Therefore, a combination of appropriate isolation and purification techniques with high resolution separation and sensitive detection methods are of importance (**Figure 2.1**).

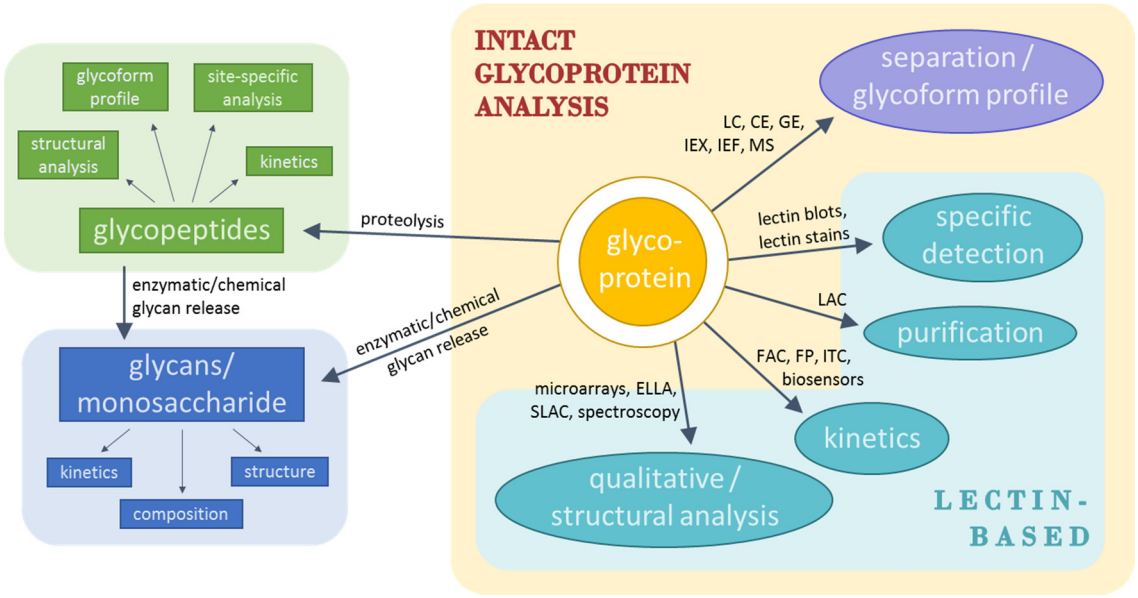


Figure 2.1 – Exemplary strategies for a detailed intact glycoprotein analysis. Direct (purple) and lectin-based (turquoise) analytical methods for glycoprotein separation, purification, detection, as well as kinetical and qualitative analysis are shown. Alternatively, proteolysis or glycan release enable further investigation of resulting glycopeptides and glycans for e.g. glycosylation site analysis or detailed structural and compositional elucidations (green and blue, respectively). Abbreviations: GE, gel electrophoresis; IEX, ion exchange chromatography; IEF, isoelectric focusing; FAC, frontal affinity chromatography; FP, fluorescence polarization; ITC, isothermal titration calorimetry; ELLA, enzyme-linked lectin assay; (S)LAC, (serial) lectin affinity chromatography.

Gel electrophoretic approaches

One approach for glycoprotein analysis is gel electrophoresis in its various forms, e.g. isoelectric focusing (IEF), native polyacrylamide gel electrophoresis (PAGE), or sodium dodecyl sulfate (SDS) PAGE [79]. During gel electrophoresis charged particles migrate in gel-forming polymers upon application of an electric field. The pore sizes of such gels can be varied in dependence on the polymer / crosslinker ratio and their concentrations [80]. Most often acrylamide and bisacrylamide are used as polymer and crosslinker, respectively. IEF gels are usually high-porosity gels with only low amounts of polyacrylamide, whereas for SDS-PAGE gels the polyacrylamide concentration is increased for sieving effects. Implementation of certain additives like the detergent SDS, the chaotrope urea, or the

reducing agent dithiotreitol (DTT) results in the denaturation of the investigated glycoproteins which is mostly required for classical SDS-PAGE. Further influences are achieved by choice of electrolyte system (e.g. tris-glycine, bis-tris, tris-acetate) and positive or negative ionic detergents. In IEF gels, proteins are separated according to their isoelectric point and, consequently, to their intrinsic charge using a pH gradient within the gel matrix. This results in a separation of the glycoprotein isoforms with high resolution and usually gives a reliable view of the degree of e.g. sialylation, sulfation, acetylation, acylation, or phosphorylation. In contrast, SDS-PAGE allows the determination of the molecular mass due to a uniform negative charge density of the (glyco)proteins caused by the covering with SDS. In the case of glycoproteins, however, the interaction of SDS with the protein can be altered by the bulky glycans leading to a decreased migration in the gel and apparently higher molecular weights [81]. Moreover, the glycoprotein bands on the gel are rather broad reflecting the huge variety of glycoforms. A combination of IEF and SDS-PAGE, the 2-dimensional PAGE (2D-PAGE), additionally reveals the mass distribution of the different isoforms. In contrast, native PAGE allows for analysis of protein aggregates, which are disrupted in SDS-PAGE under the denaturing and often reducing conditions. Yet, for studying protein aggregates, usually other methods like size-exclusion chromatography (SEC), dynamic light scattering (DLS), field-flow fractionation (FFF), and ultracentrifugation are more frequently applied [82-84].

The specificity of the described electrophoretic methods can be further increased by varying the staining of the analytes. Non-specific staining like Coomassie or silver provides an overall detection of the different glycoforms and non-glycosylated proteins. However, these techniques can lead to weak staining or even failure of detection: steric interference of the carbohydrates can hinder the binding of silver ions and an alteration of the protein hydrophobicity combined with steric hindrance by the hydrophilic glycans can result in lower binding of Coomassie, which usually interacts with hydrophobic and basic amino acids. Thus, glycan-specific methods can be used, which involve an oxidation of the glycans to aldehydes by periodic acid and subsequent staining with the Schiff's reagent (periodic acid-Schiff, PAS), Alcian Blue (a copper phthalocyanine dye), or Pro-Q Emerald dyes (fluorescent hydrazides), in the order of increasing sensitivities [85]. Still their sensitivities stay behind typical unspecific staining methods. A more sensitive specific biomolecular recognition can be performed directly with fluorescently labelled lectins or with lectins or antibodies after transferring the proteins from the gel onto membranes (lectin or immunological / Western blotting) [86, 87]. Although gel electrophoresis is a very cost-effective and widely used approach, it is on the other hand rather time-consuming, requires good technical skills for high reproducible results, and includes analyte-dependent differences in staining, which makes quantitation difficult. Molecular mass determination will show in case of glycoproteins and hydrophobic proteins a certain bias, too.

Mass spectrometric approaches

More accurate molecular weight determinations and mass distribution profiles of glycoproteins than with SDS-PAGE are achieved with MS, which has become a powerful tool in the field of glycoprotein/glycan analysis and enabled enormous progress in structural elucidation [88-90]. As with SDS-PAGE, glycoprotein MS signals are influenced by the heterogeneity of the sample. Thus, the presence of multiple glycoforms requires high mass resolution and accuracy of instrumentation to prevent over-lapping signals and, therefore, loss of information about the precise molecular composition.

The soft desorption/ionization techniques matrix-assisted laser desorption/ionization (MALDI) and electrospray ionisation (ESI) are widely used for the analysis of intact glycoproteins. MALDI scores with high sensitivity, very high molecular mass range, and a higher tolerance for contaminants, but shows drawbacks concerning the possible loss of labile monomers like sialic acid groups during desorption/ionization [91]. In order to stabilize such labile residues glycans can be derivatized. While MALDI mainly generates singly or low-charged ions, glycoproteins carry multiply charges after ESI desorption/ionisation. Consequently, data interpretation can be more challenging with the latter, especially when analysing complex samples or even not possible at all (particular in case of high content of glycans). This often requires the combination of ESI- or nano-ESI-MS with preceding separation techniques like CE or LC (e.g. SEC, ion exchange IEX, hydrophilic interaction HILIC, reversed phase RPLC). A uniform and linear response from all glycoforms represents an additional experimental challenge in MS and can be impeded by ion suppression effects or varying desorption/ionization efficiencies. Besides that, glycosylated proteins or peptides are generally more difficult to desorb and ionize [75].

Mass analysers typically used for intact glycoprotein analysis are time-of-flight (TOF), Orbitrap, Fourier-transform ion cyclotron resonance (FT-ICR), quadrupole (Q), ion trap (IT), and hybrid instruments like TOF/RTOF, LTQ (linear ion trap quadrupole), LTQ Orbitrap, or QRTOF, that combine the merits of different analysers [89]. Especially FT-ICR and Orbitrap instruments feature high mass resolution and accuracy, but require ultrahigh vacuum and long data acquisition times in return and, in the case of FT-ICR, also expensive superconducting magnets.

Next to measuring unfolded glycoproteins under denaturing conditions, native MS not only enables the analysis of non-denatured analytes and their glycosylation patterns but also of non-covalent interactions [88, 92]. Under these near-physiological conditions using volatile aqueous buffers at neutral pH, glycoproteins maintain their native structure and binding properties. Also labile groups have shown to remain attached to the glycoprotein and the sialyl residues introduce less desorption/ionisation bias. Furthermore, a combination of native MS with ion mobility spectrometry (IMS) allows for the charge and shape-selective separation of glycoforms according to their gas-phase collision cross-sections in addition to their mass spectrometric characterisation/identification [93].

For structural elucidation concerning protein sequence, localisation, and characterisation of glycosylations, intact glycoproteins can be directly subjected to MS and tandem MS analysis in a "top-down" MS-based approach. This strategy was successfully applied for the determination of O-glycoprotein sequence and O-glycosylation sites [94]. Yet, this approach is still limited to small glycoproteins in isolated form or in simple protein mixtures and the use of high performance ICR analyser.

All in all, the direct analysis of the intact glycoprotein is an attractive alternative, as it reduces sample preparation steps like enzymatic digestion or deglycosylation. With that, it speeds up the analysis, decreases sample loss and modification during preparation, and increases reproducibility.

Liquid chromatographic approaches

LC promises relative high speed, reproducibility, resolution, real-time detection, full automation, and compatibility with MS. As already mentioned above, several modes are available, which can be used for glycoprotein analysis: SEC separating analytes according to their size (hydrodynamic diameter), RPLC and HIC (hydrophobic interaction chromatography) according to their hydrophobicity, HILIC separating analytes according to their polarity, and IEX according to their charge (affinity chromatography is addressed in the next chapter). In all cases, especially when working with low analyte concentrations or doing quantifications, adsorption of the analytes to the column materials should be reduced to a minimum, as it simultaneously reduces sensitivity.

SEC can be employed as a high-throughput alternative to gel electrophoretic approaches for aggregation studies yielding information on aggregate content and size of resulting particles. Thus, it can be used for purity control and batch-to-batch consistency checks of biopolymers. It is a very robust method, yet with a low dynamic range, efficiency, and loading capacity. RPLC proved not very suited for glycoform characterization due to its lower separation efficiency for glycoforms and the general poor performance in protein separation. In contrast, HILIC and IEX have shown the ability to retain and separate hydrophilic compounds and also intact glycoproteins [74, 95]. HILIC separations according to the analyte polarity are characterized by a hydrophilic stationary phase (silica derivatized with polar functional groups like amine, amide, cyano, or diol) and a mobile phase containing high amounts of organic solvent. Retention of the analytes is influenced by hydrogen bonds, electrostatic interactions with the stationary phase, and hydrophilic partitioning between stationary and mobile phase. With its high MS-compatibility it represents a promising alternative for IEX. HIC is commonly used for (glyco)protein purifications, but only seldom for analytical purposes compared to other techniques like SEC, IEX, or HILIC [96]. Its main advantage over RP is the little denaturing effect and thus maintaining biological activity of the analyte due to non-denaturing conditions during separation.

Capillary electrophoretic approaches

Another electrophoretic approach for qualitative and quantitative glycoprotein analysis next to a slab gel is electrophoresis in a capillary. In comparison, CE offers high speed, great efficiency and utmost resolution, real-time detection, and full automation while requiring only small sample amounts. Analytes are separated according to size, shape, and physico-chemical properties allowing to resolve glycoforms. Optimizations by e.g. altering the buffer system or the pH of the background electrolyte (BGE), or by including additives like detergents or organic/inorganic molecules or ions, can further influence the resolving power. Also modifications of the capillary wall for reducing unspecific adsorption are of utmost importance, as reversible and irreversible adsorption/desorption events can decrease the separation efficiency, alter the electroosmotic flow (EOF, see **3.1.1**), and affect quantitative investigations [74, 97]. The adsorption can be reduced by varying factors like the pH and composition of the BGE, or the nature and modification of the solid surface (e.g. static or dynamic capillary coatings). Nevertheless, adsorption effects in CE is still lower than in LC methods.

Several different CE modes have shown a capability for good glycoprotein separation: capillary zone electrophoresis (CZE), capillary isoelectric focusing (CIEF), capillary gel electrophoresis (CGE), and micellar electrokinetic chromatography (MEKC) [97, 98]. CZE, however, is the most common method in routine analyses due to its rapidness and high separation efficiency, such as the separation of EPO sialoforms (glycoforms with different sialic acid contents) and identification by coupling CZE to ESI-RTOF-MS [99].

During CIEF analysis, analytes are separated according to their isoelectric point in a pH gradient formed by carrier ampholytes with very high resolution similar to IEF gels.

MEKC utilizes ionic micelles and enables the separation of also neutral analytes, which are embedded into the micelles. Depending on the distribution coefficient, the electrophoretic mobility increases with the incorporation of the analytes into the micelles [98]. The migration time is influenced by the electrophoretic velocity of the micelle and the analyte, the distribution ratio, and the velocity of the EOF.

CGE, on the other hand, resembles classic SDS-PAGE with a separation in a sieving matrix of linear or branched polymers and a reduced EOF. However, it offers some fundamental advantages to SDS-PAGE, such as improved resolution and enhanced precision in sizing, automation, speed and user-friendliness, and on-line quantitation [74]. Especially the development of miniaturized systems further decreased sample consumption and analysis time, as well as increased sensitivities and enabled routine monitoring of purity and integrity of proteins, DNA/RNA, and glycoproteins (see **3.1.2**).

Spectroscopic approaches

Spectroscopic methods can be applied to study secondary and tertiary structures of glycoproteins [74]. The latter are ideally addressed by X-ray crystallography and advanced nuclear magnetic resonance

(NMR) methods, which both require a highly purified glycoprotein. Besides that, rather high concentrations are needed for NMR studies and data analysis can be very complex. Thus, structural determination is limited to low molecular mass (< 20 kDa) glycoproteins and detailed NMR analysis of larger glycoproteins requires stable-isotope labelling [100]. For X-ray, on the other hand, (glyco)proteins have to be crystallized, which can cause difficulties for glycosylated proteins. The secondary structure can be examined with absorption and fluorescence spectroscopy, circular dichroism (CD) complemented by infrared (IR) spectroscopy. With dynamic light scattering (DLS) the molecular mass can be determined in case of more or less narrow dispersity and glycoprotein aggregates investigated.

2.2 Glycoprotein analysis using lectins

Many analytical approaches especially for the isolation and purification, but also the detection or structural analysis of glycoproteins involve the use of antibodies or lectins. In the beginning both were mostly involved in agglutination and blood typing experiments. Today, they are often applied to reduce the complexity of a sample or to specifically enrich glycoproteins in general or certain glycoforms in particular from complex biological samples. With the same approach glycoproteins can be purified for subsequent analysis when needed. Lectins and antibodies can help to detect changes in the glycosylation pattern of proteins and to specifically stain tissues or membranes. Additionally, they can be part in interaction studies for kinetic or structural elucidations.

Although monoclonal antibodies have a higher specificity and affinity towards a certain glycoprotein (i.e. an epitope), this fact also limits their range of application. They usually only recognize a specific antigen with high affinity (a particular epitope structure), which can be problematic in the case of unknown samples or very complex (in terms of glycan composition) glycoproteins. Moreover, they are more difficult and time-consuming to isolate.

In contrast, lectins are generally more stable, cheaper to produce, better characterized, and feature a variety of specificities no antibody is known for. They have the ability to bind several different ligands and, therefore, address a broader spectrum of glycoproteins. A lot of the lectins employed in glycobiology are commercially available and originate from plants, which makes supply much easier. These advantages support the choice of lectins for a comprehensive glycoprotein analysis and examples of their applications are discussed below. Next to their frequent application as tools in bioanalytical research, certain lectins have been found to act as anti-tumour, -fungal, -viral, and -insect agents [101].

Lectin affinity chromatography

For isolation and purification, but also for detection of glycoproteins with pathologically altered sugar chains lectin affinity chromatography is widely used. For this, lectins are immobilized on e.g. agarose-

(Sepharose) or silica-based stationary phase and loaded into affinity columns. Next to column-based enrichments, lectins can be covalently linked to solid carriers like magnetic agarose- or polymer beads. These beads have also been applied in automated high-throughput analyses in combination with subsequent enzymatic digestion and tandem mass spectrometry [102, 103].

For covalent coupling of the lectins, the surface is activated with different functional groups (e.g. epoxy, tosyl, carboxylic acid, **Figure 2.2**). Glycoproteins can then be specifically and selectively bound depending on the employed lectin and separated from the crude biological sample. Due to the rather low binding affinities, the interactions can be specifically disrupted and the bound proteins eluted with competitive sugar compounds. Moreover, bound glycoproteins can be sequentially eluted according to their binding affinities by using varying substances with increasing binding strengths towards the lectins. As lectin-glycan interactions are of reversible nature, lectin columns or beads can be reused for further sampling after extensive washing and regeneration. In all cases, the choice of binding, washing, and elution conditions are of utmost importance for reliable and reproducible results due to the low affinity interactions.

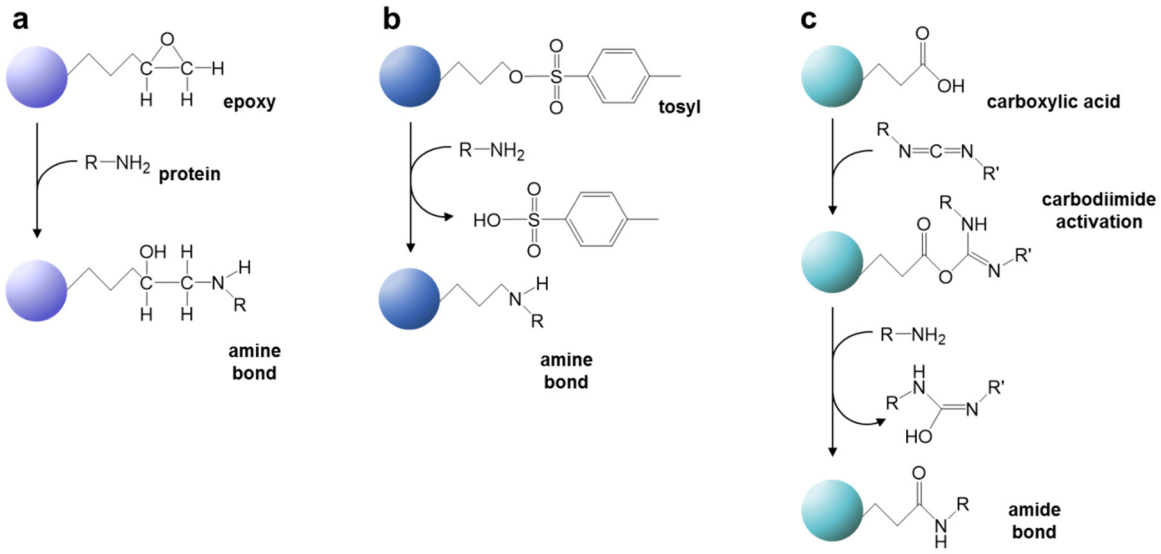


Figure 2.2 – Lectin coupling to surface activated beads. The bead surfaces are activated with various functional groups (e.g. epoxy **a**, tosyl **b**, or carboxylic acid **c**) that can be used for the covalent binding of proteins, nucleic acids, or glycans. The suitability for different ligands (as well as for glycoproteins to be enriched) can vary depending on the functional group on the bead, the related coupling conditions (stability of the ligand), the bead size, and the hydrophobicity of the bead.

Serial lectin affinity chromatography (SLAC) enables additional separation of glycoproteins in regard to differences in their glycan structures by combining columns with immobilized lectins recognizing varying carbohydrate moieties [104]. On the other hand, several lectins can be combined in one column to isolate a wide spectrum of glycoproteins from non-glycosylated species out of a

biological sample (mixed-bed or multi-lectin affinity chromatography, MLAC) [105]. In this regard, also the use of boronic acid can be considered, which binds to vicinal diols containing structures and hence to many sugar moieties [106].

Besides using lectins coupled to a solid support, glycoproteins and cells with surface-glycans can also be agglutinated and precipitated directly with lectins [34]. In the first case, the concentration of a soluble glycan is monitored that can inhibit cell-agglutination. This method particularly helped in the investigation of the human blood-groups. Precipitates, on the other hand, are the consequence of interactions between the multivalent lectin with multivalent glycoproteins forming insoluble cross-linked complexes. Titration experiments are usually used to monitor the precipitation.

Exemplary methods for kinetic and structural analysis

Kinetic interaction studies can be performed with frontal affinity chromatography (FAC), isothermal titration calorimetry (ITC), or lectin biosensors. Additional information about structural specificities can be gained with enzyme-linked lectin assays (ELLAs), which are highly similar to ELISAs (enzyme-linked immunosorbent assays), and lectin microarrays.

FAC as a special form of affinity chromatography allows kinetic investigations of affinity interactions [107]. Characteristic for this method is the continuous injection of the ligand to an immobilized binding partner. This binding partner is either bound chemically or by affinity to the stationary phase. By monitoring the glycoconjugate concentration of the elution front with e.g. UV spectrophotometry, MS, or fluorescence spectroscopy, equilibrium dissociation and binding constants can be determined [108-110]. In this regard, ligands showing no affinity towards the immobilized binding partner elute first, whereas the elution of interacting ligands is retarded. FAC is particularly suitable for investigating weak interactions as in the case of lectin and glycans and analysis is concentration-independent (detection of ligands at concentrations significantly below their respective equilibrium dissociation constants) [108]. Furthermore, miniaturization of the technique enables high sensitivities and low sample consumption as well as automation allows finally high-throughput analyses. However, interactions studies can be influenced by the immobilization of one binding partner, whose binding activities and specificities may be altered. Next to that, the complexity of samples can be a limiting factor: due to the continuous-infusion, the interacting ligand must be detected in the presence of non-binding analytes. Additional fractionation before detection can improve the compound screening rate while at the same time reducing the compound concentrations.

With ITC, interactions can be completely characterized concerning thermodynamics and kinetics of the binding [111, 112]. However, high sample amounts are needed for measurements. ITC determines the interaction induced heat difference relative to a reference cell (containing water or buffer) during titration of one binding partner to the other within an adiabatic jacket. The heat change is directly related to the enthalpy of reaction. With increasing concentration of the added analyte, the

change in heat difference minimizes due to saturation of the binding partner in the reaction cell resulting in a binding curve and thermodynamic properties (binding affinity, enthalpy, entropy, and stoichiometry) of the interaction. Despite the high sample consumption, the analytes are unlabelled and in solution, which enables uninfluenced kinetic studies. Moreover, ITC scores with ease-of-use, rapidness, accuracy, and is compatible with a wide range of buffers and solutions. Thus it has become an invaluable tool in many branches of science from cell biology and enzyme kinetics to food chemistry and lectin-glycoprotein studies [113, 114].

Biosensors also allow the kinetic investigation of lectin-glycoprotein interactions and quantitative conclusions in real-time [115]. One can differentiate between electrochemical, optical, piezoelectric, and thermal devices among others with regard to the transducer element. Typical examples are surface plasmon resonance (SPR), quartz crystal microbalance (QCM), surface acoustic wave (SAW), electrochemical impedance spectroscopy (EIS). In all cases, one interaction partner is immobilized to the sensor surface and binding is recognized by changes in the respective signal. SPR as an optical instrument, for instance, monitors the refractive index of the interface between two different media (e.g. gold or aluminium and a dielectric) [116], whereas QCM and SAW (piezoelectric) use bulk or surface acoustic waves [117, 118]. EIS measures the difference in impedance of a constant applied current at the surface of an electrode such as gold or platinum [119]. Ligands are usually immobilized via amine, thiol, or aldehyde functional groups or by streptavidin/biotin affinity on self-assembled monolayers (SAMs). Although those methods are label-free, the binding partner recognition can still be affected by the immobilization. Nevertheless, biosensors offer highly sensitive and rapid kinetic and quantitative information and can be found in a variety of fields including clinical diagnosis, environmental and military monitoring, genetic screening, and pharmaceutical applications.

Like ELISAs , ELLAs are performed in microtiter plates and lectins are used for the detection of interactions with glycoproteins or glycan structures found on cell surfaces [120]. As with biosensors, one of the binding partners is immobilized, but in this case by non-specific adsorption. The other binding partner is usually either directly labelled with alkaline phosphatase, peroxidase, or a fluorescence dye for detection or it is biotinylated and a labelled streptavidin or antibody is used. However, labelling of the analytes can influence the binding properties. Alternatively, a sandwich assay format avoids the need for direct labelling of the interaction partners [121, 122]. Here, glycoproteins are immobilized by deglycosylated antibodies and a labelled secondary antibody or lectin is used for detection of the bound lectins. The amount of bound lectins is then determined with automated ELISA plate readers or by fluorescence. Quantitative evaluations during ELLAs can be impeded by a non-uniform orientation of the immobilized binding partner, varying coating densities, and non-specific binding. Nevertheless, ELLA is a very rapid and sensitive method to gain insights in structural specificities.

A comparably new method for studying binding specificities was developed in 2005 with lectin microarrays [69, 123, 124]. For analysis of glycoproteins, but also cells, bacteria, or crude extracts, a series of well-characterized lectins is printed in spatially separated spots onto a solid surface (multi-well plates, microarray chips, or 3D hydrogels). Immobilization can be of covalent nature or by affinity. An additionally controlled orientation of the lectins during immobilization can increase the sensitivity of the array. In order to decrease batch-to-batch variations of naturally isolated plant lectins due to their own glycosylation, deglycosylated lectins or bacterial recombinant lectins are preferred. Usually, the analytes bound to the immobilized lectins are detected with fluorescence either by directly labelling the analytes or by using labelled probes (e.g. biotinylated antibodies against the glycoprotein and fluorescently labelled streptavidin). The lectin microarrays allow rapid and high-throughput investigations with only low sample consumption and preparation. Sample purification is not necessarily required and also glycans do not have to be released for a better analysis. Due to the huge variety of possible samples they have been extensively used in medical and therapeutical applications. They could contribute to the research of cancer, stem cell, bacteria, fungi, diabetes or other diseases with valuable qualitative, structural information.

Nowadays, a wide array of different analytical methods is available for a detailed study of glycoproteins. Most of the times, lectins can contribute in many different ways supporting purification, isolation, detection, as well as qualitative and quantitative elucidations. The choice of method can vary depending on the information needed and based on the sample available (amount, crude or purified). Also the influences created by labelling or immobilization of the analytes have to be taken into account. Moreover, the time-factor and the need for high-throughput can exclude certain techniques. Often, a combination of various methods is beneficial for a general and overall analysis. Thus, high sensitivity, specificity, and reliability, as well as low sample consumption and rapidness are preferred, which encourages a constant search for improvements of analytical methods.

3. Alternative approaches for glycoprotein and glycoprotein-lectin analysis

3.1 Liquid-phase analysis: microchip capillary gel electrophoresis

Despite its advantages like high separation efficiency, low sample consumption, automation etc., CE took much longer to become a standard tool in analytics and catch up to other techniques as different LC modes or GE. One reason is simply its comparable late introduction in the early 1980s [125]. Moreover, it stayed behind in preparative-scale separations and analyte recovery [126]. However, especially its high resolving power has turned CE into an important asset in the fields of diagnostic, pharmaceutical, and industrial research. New developments in microscale techniques additionally rose interest for rapid and easy to handle high-throughput CE-on-a-chip analyses.

3.1.1 Principles of CE

In general, electrophoretic separations can be performed with or without a supporting medium. In contrast to electrophoresis in a gel as solid support, CE generally uses narrow bore fused-silica capillaries filled with BGE as channels for electromigration. Due to their typical dimensions of 20-100 µm inner diameter and 20-100 cm length they allow for efficient dissipation of Joule heat generated from the applied field. The resulting possibility of using higher voltages enables higher separation speeds for CE compared to gels. Furthermore, these dimensions also allow the use of only low amounts of buffer and sample introduced into the capillary with volumes in the µL to nL range, respectively.

Based on CZE, the basic and most commonly applied mode of CE, the other forms are usually derived from it by addition of linear or cross-linked polymers (CGE), surfactants (MEKC), or ampholytes (CIEF) to the BGE. In the case of CZE, ionic analytes are separated in a capillary filled with an appropriate BGE at a defined pH before applying an electric field. The migration of the analytes towards the detector is determined by influences of electroosmotic and electrophoretic forces and leads to their separation according to charge, shape, and size [98, 127].

Their electrophoretic mobility µ depends on the charge *q* and frictional forces, which are strongly influenced by analyte shape and size. µ can be described by the Debye-Hueckel-Henry theory **(1)** and can be related to the mass of the particle via the Stokes' radius *r* in the case of spherical particles.

$$\mu = q/6\pi\eta r \qquad \textbf{(1)}$$

η: buffer viscosity

µ is the analyte specific factor of proportionality between the migration velocity *v* of the analyte and the electric field strength *E* (**2**). A constant velocity of the charged analyte upon application of an

electric field results from the balance (**3**) between the electric force F_{el} (**4**), that accelerates the analyte, and the frictional force F_{fr} (**5**), that retards the analyte migration.

$$v = \mu E \qquad \textbf{(2)}$$

$$F_{el} = -F_{fr} \qquad \textbf{(3)}$$

$$F_{el} = qE \qquad \textbf{(4)}$$

$$F_{fr} = -f_c v = -6\pi\eta r v \qquad \textbf{(5)}$$

f_c: friction coefficient

For weak acids or bases like proteins the mobility is dependent on their degree of dissociation α, which leads to the effective mobility μ_{eff} (**6**).

$$\mu_{eff} = \sum \alpha_i \mu_i \qquad \textbf{(6)}$$

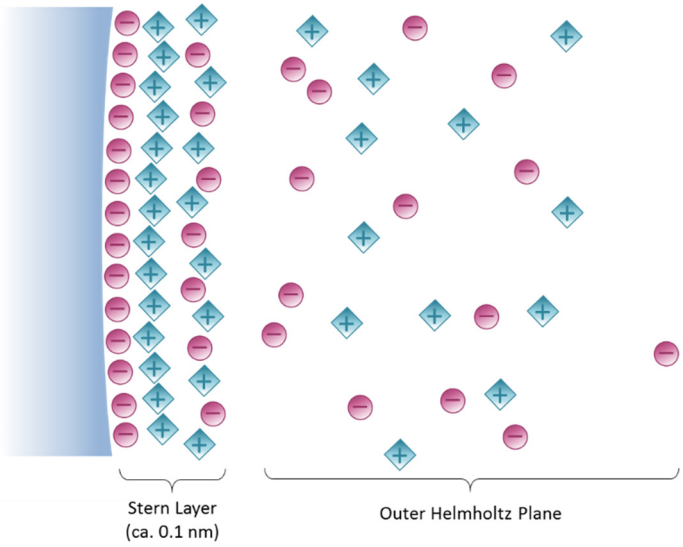


Figure 3.1 – Electroosmotic flow (EOF). The negatively charged silanol groups of the uncoated capillary wall attract positively charged ions from the buffer and cause the formation of the static *Stern Layer*. In contrast, the cations of the second Outer Helmholtz Plane start to migrate upon application of an electric filed and thus introduce a bulk movement inside the whole capillary. (according to [98])

Next to electrophoretic forces, the migration of the analytes is influenced by electroosmosis. This effect is the result of the formation of an electrostatic bilayer at the capillary walls (**Figure 3.1**). At higher pH (already > 2) the silanol groups of the uncoated fused-silica walls are starting to be deprotonated and attract hydrated cations from the BGE. The formed double layer closest to the wall is more or less static and is called *Inner Helmholtz* or *Stern layer*. Additionally, a more diffuse second

layer (*Outer Helmholtz Plane*) is formed more distant from the wall. Upon application of an electric field, the cations of this external layer start to migrate towards the cathode, while dragging along water from hydration. This introduces a movement towards the cathode inside the whole capillary, which is termed electroosmotic flow (EOF). Hereby, the strength of the EOF is positively related to the pH and inversely to the ionic strength of the BGE.

The EOF strongly affects the separation of the analytes: if the EOF is higher than the electrophoretic force acting on an anion, which is electrostatically attracted towards the anode, it is still migrating towards the cathode together with the EOF (**Figure 3.2**). Therefore, the EOF is an important value for separation that can be controlled and adapted (e.g. by varying the BGE, surface-coating of the capillary, additives, etc.).

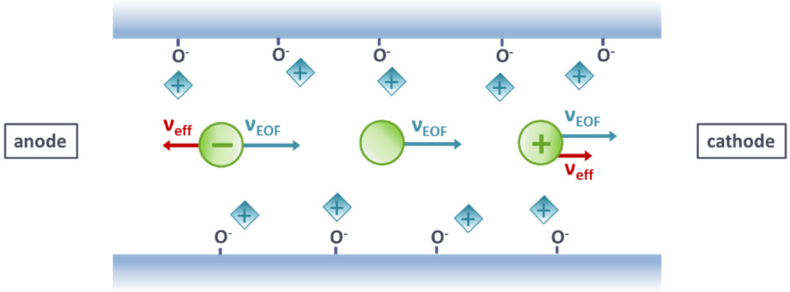


Figure 3.2 – Migration of analytes. The migration velocity of analytes is influenced by electrophoretic and electroosmotic forces. The EOF ultimately allows for the migration of also neutral analytes and analytes similarly charged as the detector electrode (cathode).

Taking the EOF into account leads to an analyte velocity *v* as defined by **(7)**, which contains the mobilities due to the electric potential μ_{eff} and the EOF μ_{EOF}. The resulting apparent solute mobility μ_{app} depends on external variables like temperature and solvent (e.g. pH, ionic strength, viscosity) as well as analyte specific properties like particle size, shape, and charge.

$$v = \mu_{app}E = (\mu_{eff} + \mu_{EOF})E \qquad \textbf{(7)}$$

The analytes are most frequently detected by on-capillary UV/Vis or fluorescence measurements. UV-Vis detectors make use of natural light absorption of proteins. Peptide bonds and aromatic amino acids can absorb electromagnetic energy at 200-220 and 260-280 nm. As this detection often lacks sensitivity, laser-induced fluorescence (LIF) detection is a good alternative. Here, proteins are labelled with a fluorophore, which is excited by a laser (solid-state lasers, gas lasers, semi-conductor/diode lasers, or light emission diodes). The emitted fluorescence can then be measured. Many different fluorescence dyes have been developed and studied spanning a huge range of different excitation and emission wavelengths [128, 129].

Fluorophores can be differentiated into fluorogenic, which only fluoresce when bound to their target, and fluorescent, which fluoresce whenever being excited [130]. The latter suffer from higher background fluorescence and thus lower sensitivity, when applied for protein detection, and often require prepurification from the unreacted dye. For derivatization many different labelling strategies for various target groups can be applied. Next to carboxyl, carbonyl, hydroxyl, chloroalkyl, reducing aldehyde/keto, or sulfhydryl groups, amines, especially lysine residues as one of the most abundant amino acids, are most commonly used as derivatization sites. However, their occurrence can vary between analytes and influence the labelling efficiencies. Additionally, derivatization with neutral or charged dyes can alter the overall charge of the (glyco)protein and, therefore, influence its electrophoretic properties. Besides, often high (glyco)protein concentrations are needed for covalent labelling and the process itself is very time-consuming requiring additional prepurification steps. Moreover, incomplete or unequal labelling (e.g. due to steric hindrances and accessibility) can create multiple different species with various numbers of dyes attached, which migrate independently impeding sizing and quantitation [131]. To overcome this problem, the use of a minimal labelling strategy ensures that only one single target residue is labelled in each protein. Therefore, the ratio between dye and protein is kept deliberately low, which results in the derivatization of less than 3 % of all proteins.

In the case of CGE, migration is as well influenced by electrophoretic and electroosmotic forces. However, due to the presence of sieving matrices, the frictional forces are more influential. Moreover, depending on the polymer concentration (> 4 %), electroosmosis is reduced significantly without a displacement of the gel during the run even when no capillary wall coatings are used [132]. This fact can be attributed to the higher viscosity or to polymer adsorption onto the capillary wall at higher concentrations, which additionally stabilizes the gel. As a consequence of EOF reduction and increase of frictional forces, proteins are mainly separated according to their size and shape. The contribution of the charge to the mobility is suppressed by the addition of SDS, which enables quantification together with molecular weight determination by comparison to a calibration curve of known standards. Compared to SDS-PAGE the time-savings in CGE are relatively low. The down-sizing of CGE from lab-scale to chip-scale (e.g. with the Agilent 2100 Bioanalyzer, the Caliper LabChip90, the BioRad Experion, or the Shimadzu MCE-202 MultiNA), on the other hand, introduced several new advantages over SDS-PAGE.

3.1.2 Microchip capillary gel electrophoresis of proteins with the Agilent 2100 Bioanalyzer

The Agilent 2100 Bioanalyzer system was introduced in collaboration with Caliper Life Sciences in 1999 as the first chip-based capillary gel electrophoretic instrument capable of analysing proteins next to DNA, RNA, as well as cells in simple flow cytometry experiments (**Figure 3.3**) [133]. By using a

microfluidic system of microchannels without any moving parts it combines sample handling, separation, staining, and detection on one chip [134-136]. The Lab-on-a-chip enables a rapid, reliable, and sensitive separation and quantitation of analytes with LIF detection.

The applied microfluidic chip is made from two soda-lime glass layers thermally bonded together, cut to 17.5 x 17.5 mm square, and implemented onto a polymer encasement. The first layer contains the microchannel system (13 and 36 µm wide), which are fabricated by photolithography with a positive photoresist and chemical wet-etching [137, 138]. The second layer comprises 16 holes for buffer and sample reservoirs and provides access to the microchannels (**Figure 3.4**).

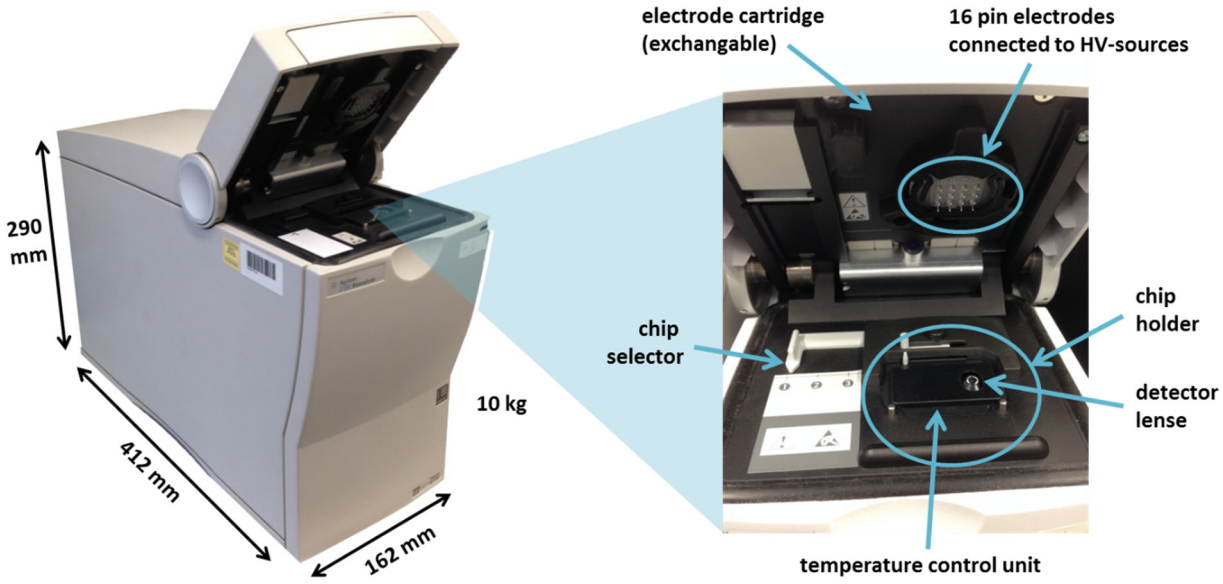


Figure 3.3 – The Agilent 2100 Bioanalyzer system. The table-top instrument contains an electrode cartridge for the analysis of proteins and nucleic acids, which can be exchanged for flow cytometric measurements. The chip holder positions the chip for LIF detection and includes a temperature control unit for investigations at constant temperature (30 °C in the case of protein chips). The setting of the chip selector depends on the type of chip used (protein, DNA, or RNA chip). Each of the 16 electrodes is dipped in one well for individual application of voltages.

For analysis, the channels are filled with a sieving matrix consisting of a linear polymer forming dynamic pores directly before measurement. As the EOF is efficiently reduced by this matrix, analytes on the chip are moved by electrophoresis only. 16 independent platinum electrodes connected to high-voltage power supplies can be programmed individually for each well to induce a controlled movement in speed and direction. In this way, the analytes are transported towards the 1.25 cm long separation channel. A small sample plug of approximately 25 pL is injected by a switch of the electric fields using a two-step injection process (**Figure 3.5**). In less than a minute the analytes in the sample plug are then separated according to their size in the SDS containing sieving matrix and detected by LIF. For that the

instrument is equipped with a red laser diode (635 nm excitation, 685 nm emission) and/or a blue LED (475 nm excitation, 525 nm emission, mainly used for flow-cytometry). In order to increase accuracy and reproducibility of sizing, internal standards are added to each sample to allow for corrections of small migration drifts during the course of a chip run. Finally, data are presented in classical electropherograms, but can also be displayed as gel-like images for better understanding and comparability to SDS-PAGE results.

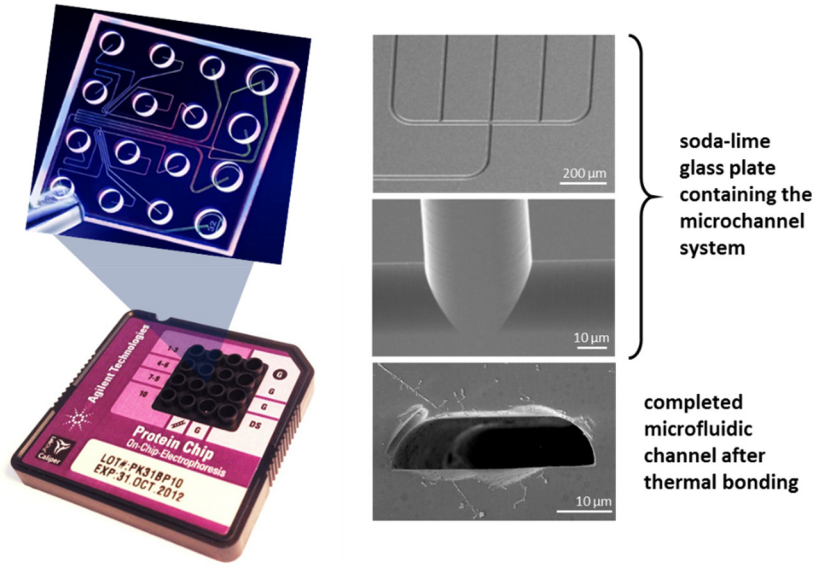


Figure 3.4 – The design of a Bioanalyzer microchip. The microfluidic system is formed by two thermally bonded soda-lime glass plates using photolithography and chemical wet-etching. The chip is then mounted onto a plastic caddy resulting in the commercially available LabChip with 16 wells. (Microscopic pictures according to the Agilent E-Seminar „Agilent 2100 Bioanalyzer: One Platform – Endless Possibilities“; http://brunell.org/advbiotech/01-23%20Electrophoresis%20and%20Bioanalyzer.pdf, 27.08.2016)

LIF detection requires the labelling of the analytes with fluorescence dyes. The available protein assays for the Bioanalyzer offer two possibilities (**Figure 3.6**): covalent analyte labelling (High Sensitivity Protein 250 / HSP-250 Kit) and non-covalent labelling by binding of the dye to SDS-analyte complexes (Protein 230 or 80 / P230 or P80 Kit). The latter is a dynamic labelling strategy performed on-chip by mixing the dye with the sieving matrix before introduction into the microchannels. The analytes are non-covalently labelled in less than 100 ms, when getting in contact with the dye-sieving matrix mixture in the separation channel (**Figure 3.5**). As this dye binds to SDS-protein complexes as well as to SDS micelles, this staining method would generate a high background fluorescence, which would decrease the sensitivity of the assay [139]. This effect is circumvented by the dilution of the SDS concentration below the critical micelle concentration and a resulting removal of the “empty” SDS micelles. As a consequence, the freed dye can additionally bind to the SDS-protein complexes. Dye,

not included in any micelles or SDS-protein complexes shows significantly reduced fluorescence. Therefore, the background signal is not only reduced, but the signal intensity increased at the same time by one order of magnitude [134]. For instrumental realization an extra intersection with two channels confining the separation channel is included on the chip before detection (**Figure 3.6a**). These channels contain the sieving matrix, yet no dye and SDS. The sensitivity of this method can be compared to standard Coomassie Brilliant Blue stained SDS-PAGE gels (10 – 100 ng [85]) and has a linear dynamic range over two orders of magnitude, which is usually smaller in SDS-PAGE analyses [135]. In contrast to SDS-PAGE, however, the sensitivity can also vary depending on the ionic strength of the analyte sample buffer.

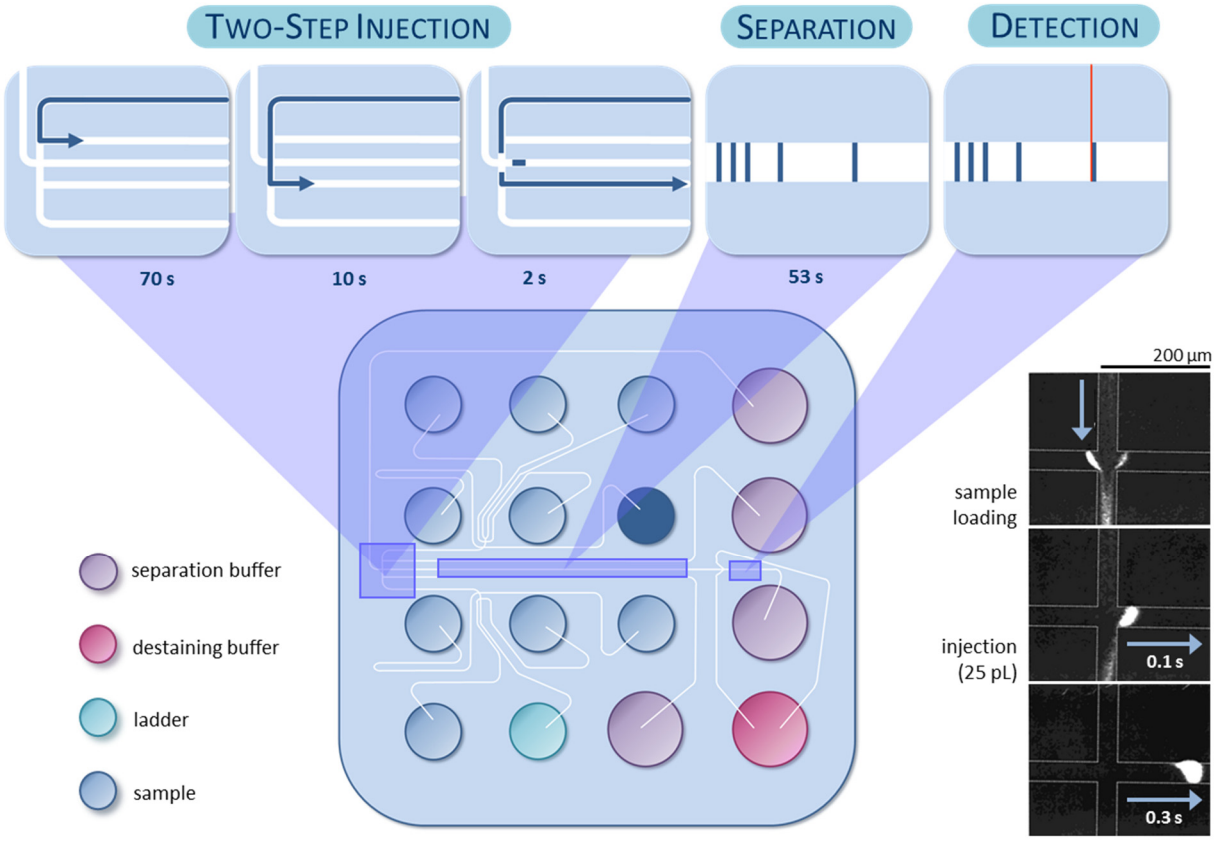


Figure 3.5 – The analysis of a sample on a Bioanalyzer microchip (on-chip staining). Upon application of an electric field the sample (dark blue) is electrophoretically moved towards the separation channel. By a switch of the electric fields, a small sample plug (ca. 25 pL) is introduced into the separation channel and analytes are stained within 0.1 s (see microscopic pictures). Analytes are separated according to their size and detected by laser-induced fluorescence (LIF) with a red laser diode in the case of proteins. (Chip layout and microscopic pictures according to [134])

Alternatively, the analytes can be covalently labelled with fluorescence dye before chip analysis (**Figure 3.6b**) [140]. The respective assay uses a dye that introduces one negative charge after labelling

and binds to the ε-amino-groups of Lys by a N-hydroxy succinimidyl (NHS) ester. In order to guarantee the most comparable labelling efficiency, a minimal labelling procedure is applied. Thus, only about 3 % of the proteins are tagged with statistically one dye molecule. Excess reactive dye is quenched with ethanolamine and not removed from the sample before analysis, as it is co-migrating with the internal standard. The method showed an extended linear dynamic molecular range of four orders of magnitude and sensitivities comparable to silver stained SDS-PAGE gels with a minimal protein amount of 0.5 ng required for labelling (LOD for silver stained gels: 0.2 – 10 ng) [140]. Other assay characteristics concerning sizing and quantitation were comparable to kits using the on-chip staining method.

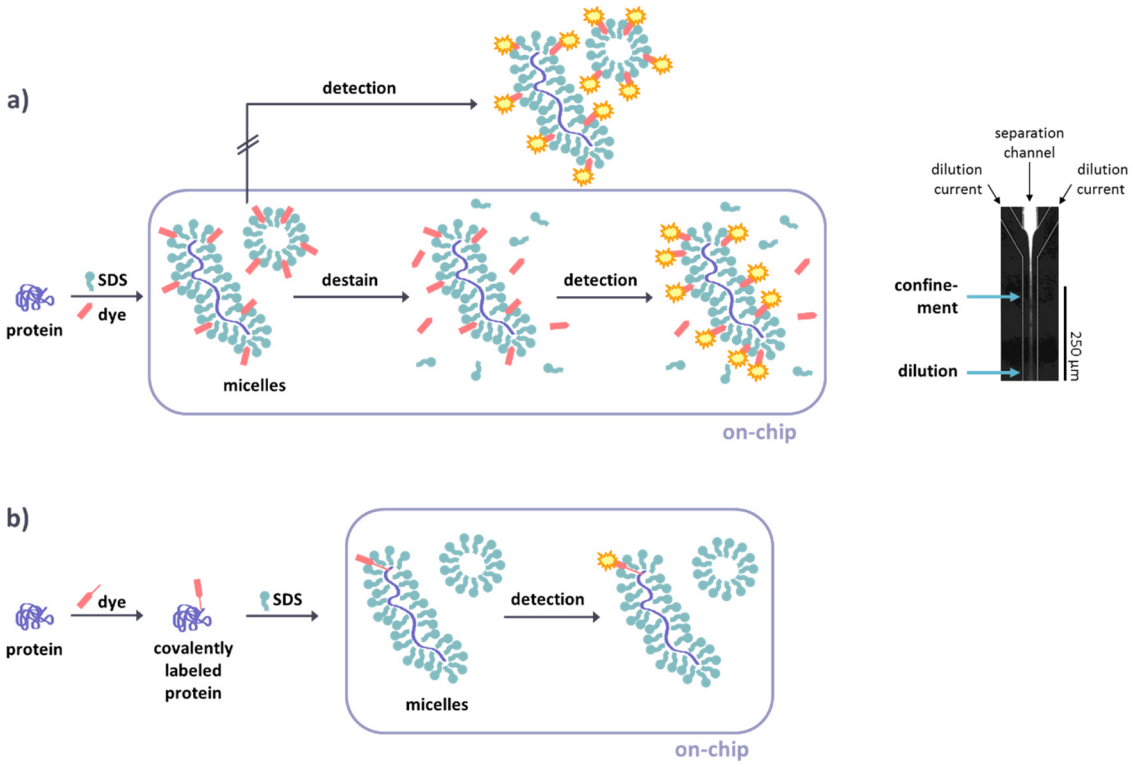


Figure 3.6 – Methods of staining on 2100 Bioanalyzer protein chips. During dynamic on-chip staining a non-covalent dye intercalates into a SDS-protein complex (**a**). In order to reduce high background caused by dye binding to SDS micelles, the SDS concentration is diluted below the critical micelle concentration (cmc) resulting in a break-up of the SDS micelles. An intersection with two channels adjacent to the separation channel is introduced post-column before detection, whose electrical currents confine the SDS to a thin stream and finally induce dilution (see microscopic picture according to [134]). Covalent fluorescence labelling requires a pre-chip staining step (**b**), but allows for higher sensitivity and linear range. A minimal labelling strategy reduces side-products with more than one dye attached and thus increases accuracy of quantitation.

Sizing of analytes is realised by comparison to standards of known molecular weights. The sizing accuracy strongly depends on the properties of the investigated analytes (isoelectric point, structure, modifications etc.), which is also observable for SDS-PAGE. However, the Bioanalyzer assays have

shown higher reproducibility and accuracy compared to the latter for standard proteins as well as real biological samples like serum proteins [135, 141]. Moreover, quantitation accuracy and reproducibility, which are determined by staining efficiency and again analyte properties, were determined to be better with those protein assays. Next to those advantages, Bioanalyzer assays are much faster, easier to handle, and more reliable. They consume less sample and buffers, produce less waste, and are less hazardous. The Bioanalyzer system enables the analyses of RNA, DNA, and (glyco)proteins combined in one instrument. Furthermore, due to the high automation they are perfectly suitable for clinical studies and quality control of biopharmaceuticals and the derived biosimilars.

Although the assays come as pre-prepared kits, the possibility to individually introduce software modifications to the Bioanalyzer script due to a close scientific collaboration with Agilent enabled adjustments for other applications. Thus, a chip-based CE using the influence of the EOF (CE-on-a-chip) and without a sieving matrix could be developed, which required a change of polarity [142]. This allowed for the separation of large particles like viruses and protein-complexes according to charge and size on chips. On the other hand, the matrix composition of the commercially available kits is optimized for given size-ranges and can to a certain extent be adapted to achieve a higher resolving power or extended size range.

Due to the recent interest in the study of post-translationally modified proteins, especially glycoproteins, also the Agilent 2100 Bioanalyzer was considered as a promising new tool for sensitive and reliable investigations. However, the commercially available protein assays are mainly designed for low or completely unmodified proteins. Only little and contradictory data was available in literature before the beginning of this work in regard to the applicability of the Bioanalyzer system for glycoprotein analysis [143, 144]. This made a decent assay characterization in regard to sizing, quantitation, reproducibility, and sensitivity necessary and of general interest, before application of this device to the analysis of unknown or complex biological samples (further details see **Publications I** and **II**).

3.2 Gas-phase analysis: nano electrospray gas-phase electrophoretic mobility molecular analyzer

Next to investigations in liquid-phase, the nano electrospray gas-phase electrophoretic mobility molecular analyzer (nES GEMMA) separates analytes in gas-phase according to their electrophoretic mobility diameters (EMDs), which can be directly correlated to the particle diameters in the case of spherical particles [145-148]. With nES GEMMA characterizations being generally independent of the analysed particle type, i.e. chemical nature of analytes of interest, the system has proven a broad applicability for analytes ranging from only a few nm in size up to several hundred nm. Examples can be found for DNA [149], proteins and glycoproteins [145, 146, 148, 150], polymers [151-153], viruses and virus-like-particles [150, 154-156], or gold nanoparticles [157, 158]. Additionally, nES GEMMA has

shown its strength to preserve even fragile non-covalent interactions, as it is operated at ambient pressure and with non-denaturing electrolyte solutions [159-163]. This fact would render it a new and very suitable tool for analysing the rather weak glycoprotein-lectin interactions directly from solution, in which they have formed.

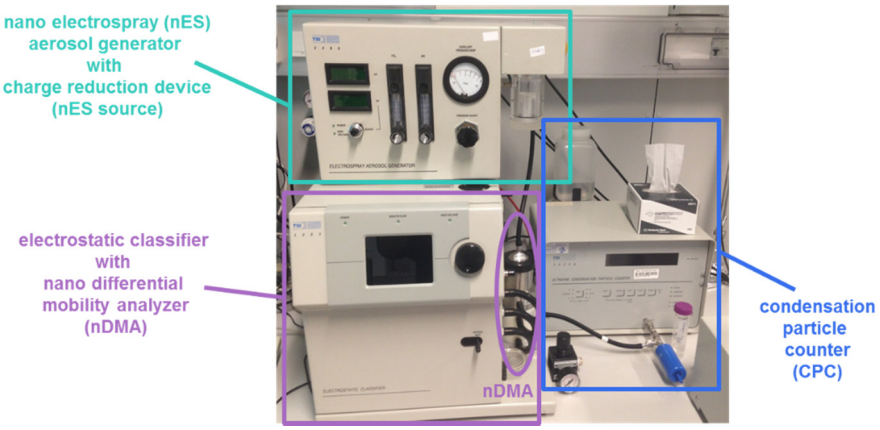


Figure 3.7 – Set-up of the employed nES GEMMA system. The instrument consists of a nES source with a ^{210}Po source for charge reduction, a nano differential mobility analyzer (nDMA) for analyte separation, and a condensation particle counter (CPC) for detection.

For nES GEMMA analysis, multiply charged droplets are produced from the liquid sample in the nES unit operating in cone-jet mode (**Figure 3.7**). By drying and a simultaneous charge reduction in a bipolar atmosphere, mainly neutral and singly charged particles are created. Only the charged analytes with a certain EMD can exit the nano differential mobility analyser (nDMA) at a particular applied voltage. The possibility to adjust the voltage enables the scanning of a certain size range (a few nm to hundreds of nm). The obtained monodisperse aerosol is directed through a condensation particle counter (CPC), in which supersaturated n-butanol vapour condenses onto the particles. Following nucleation, single particles can be counted by laser light scattering yielding a number-based particle concentration. In the resulting diagram the EMDs of the analytes are plotted against the particle counts per volume. The molecular weights of analytes can finally be calculated by applying a correlation derived from standards of the appropriate compound class [148].

3.2.1 Nanoelectrospray generation and charge reduction

For the generation of the nES, analytes in a volatile electrolyte buffer are transported through a fused silica capillary using pressure and an electric field (**Figure 3.8**). The flow rate is dependent on the inner diameter and the length of the capillary, the applied pressure and electric field, and the viscosity of the employed buffer. By application of an electric field (between a Pt electrode in the sample and the

orifice plate as counter electrode), surface charges are induced in the liquid leading to the formation of a characteristic Taylor cone [164, 165]. An accumulation of similar charged ions at the capillary tip results in the emission of multiply charged primary droplets. Working in the cone-jet mode, very small and uniform droplets are released from a jet formed at the tip of the cone. The spraying process can be influenced by many parameters like properties of the sprayed liquid (e.g. conductivity, viscosity, surface tension), of the capillary (e.g. flow rate, capillary dimensions), or of the employed sheath gas (e.g. flow rate, temperature).

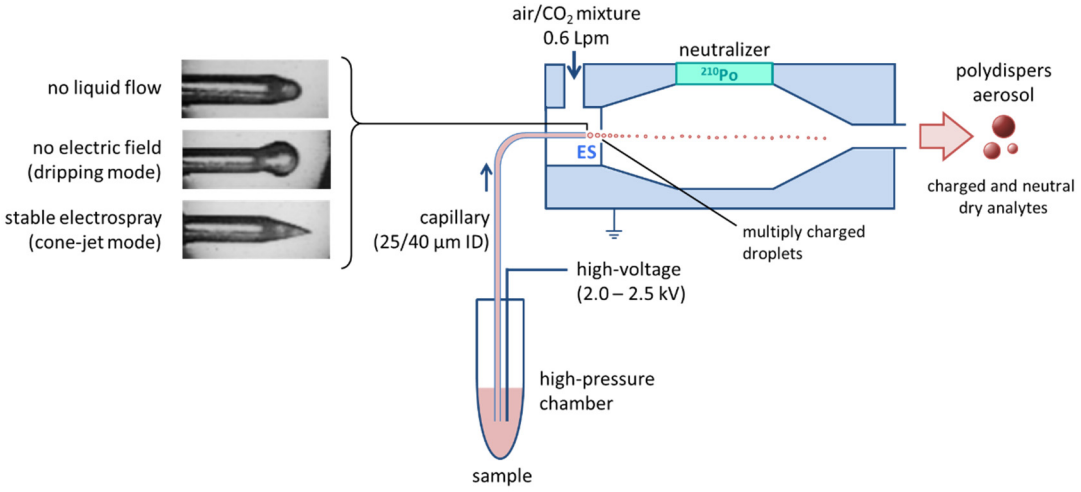


Figure 3.8 – Exemplary nES source with charge reduction device. Analytes are electrosprayed from a volatile electrolyte solution through a cone-tipped fused silica capillary. In order to generate a stable electrospray the nES unit is operated in cone-jet mode by variation of the electric field applied to the tip of the capillary. The resulting multiply charged droplets are led into the neutralizer where they are dried and simultaneously charge reduced in a bipolar atmosphere induced by an α-particle emitter (^{210}Po source). A polydisperse aerosol of mainly neutrally, a small amount of singly, and a negligible amount of multiply charged particles is created. (according to [166])

Supported by a sheath gas flow of dried and particle-free air and CO_2 the charged droplets are dried and simultaneously led into the charge reduction device. A mere shrinking of the droplets by evaporation would increase the surface charge density and the Coulombic repulsive forces, which ultimately results in Rayleigh disintegrations and a less monodisperse aerosol. Thus, the charges of the droplets are reduced in a bipolar atmosphere created by a α- or β-radiaton (^{85}Kr, ^{241}Am, or ^{210}Po). By collisions with gaseous ions (e.g. H^+, H_3O^+, $(H_2O)_nH_3O^+$, N_2^+ , O_2^+, NO^+, O^-, O_2^-, NO_2^-) the charges of the particles are steadily decreased resulting in a well-described charge-distribution (Boltzmann distribution) dependent on the particle size [167, 168]. As a consequence, most analytes are neutrally charged and only a small, size-dependent amount is singly charged (< 5 % for particles below 10 nm). The ratio of multiply charged particles is almost negligible (<< 0.01 % for particles below 20 nm).

Alternatively to the radiation-based bipolar chargers, corona dischargers (monopolar) and soft X-ray radiation discharger (bipolar) have been employed for charge reduction [169, 170].

The evolvement of the small charged droplets into solvent-free ions by evaporation is a highly discussed subject with several theories trying to describe the electrospray process (mainly for the process as found in ESI-MS): Among them are the ion evaporation model (IEM) [171, 172], the chain ejection model (CEM) [173, 174], and the charged-residue model (CRM) [175]. The latter proposes that droplets in the size of nanometres, containing only one analyte molecule, are created due to solvent evaporation and disintegration processes. This model applies more to large globular species, whereas the behaviour of low molecular weight analytes are more described by IEM [176]. IEM suggests that the droplets shrink by evaporation until the field strength at their surface is sufficiently large that the ions in their solvated states can be expelled from the droplets. The CEM, on the other hand, concerns mostly large, unfolded proteins. Due to their unfolded state, nonpolar residues, which are usually buried within the hydrophobic core of a globular protein, are now solvent accessible. This unfavourable state of a hydrophobic protein residing within the aqueous droplet interior causes a migration to the droplet surface, the expulsion of a chain terminus, and a stepwise sequential ejection of the whole protein.

In conformance with the CRM, non-volatile impurities in the sample (e.g. buffer constituents) can be detectable during GEMMA analysis, as they dry as a crust around the analyte and cause a shift in the observed EMD [146, 177]. In the same manner artificial analyte di- to oligomers can be formed and detected in case of too high analyte concentration. At low analyte concentrations the majority of the droplets are empty. However, rising concentrations increase the possibility for the occurrence of two or three analytes per droplet [178, 179]. After evaporation, they can be observed as aggregates, i.e. dimer, trimer etc., which has to be considered in the case of biological interaction or aggregation studies. These species are directly dependent on the analyte concentrations and disappear in case of reduced analyte concentrations whereas true specific aggregates or biospecific di- to oligomers can be still detected.

3.2.2 Differential mobility analysis of analytes

The DMA is a special form of ion mobility spectrometry (IMS) originally developed for the classification of gas-suspended micrometre to sub-micrometre particles and a long established method in aerosol physics and environmental analyses [180]. With instrumental developments and optimizations towards the lower nm range the technique also found attention in the chemical and biochemical field [145, 146, 148]. In contrast to more traditional IMS, DMAs employ a combination of electrical field and a perpendicular sheath flow, in which the ions are separated by their electrophoretic mobility rather in space than in time.

The particles coming from the nES source enter the cylindrical nDMA and flow together with the inner core sheath air between the electrical grounded outer electrode and the inner electrode, to which a tuneable voltage is applied (**Figure 3.9**). The voltage can either be negative or positive depending on the favoured charge to be detected. In the case of a negative voltage, only positively charged analytes are attracted towards the inner electrode, whereas negative particles are repelled and neutral ones leave the nDMA unaffected with the sheath flow. As a result, the positively charged particles precipitate on the centre electrode and merely particles with a certain electrophoretic mobility or EMD at a specific applied voltage can exit the nDMA towards the detector at an inner opening. Positively charged particles with a too low electrophoretic mobility cannot pass the sheath flow in time and thus do not collide with the electrode or hit the opening but move with sheath flow to outlet (lower part) of the nDMA.

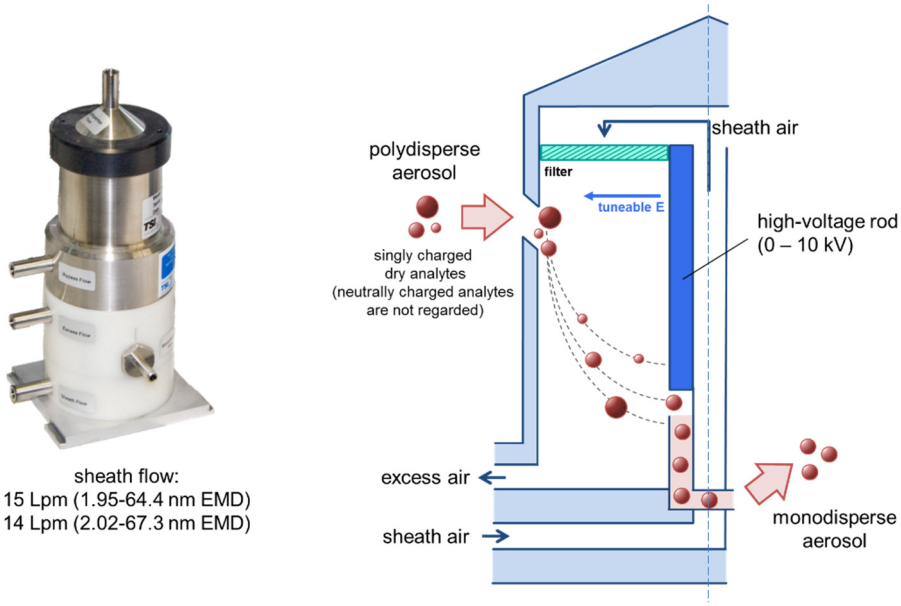


Figure 3.9 – Principle of electrophoretic separation in a nDMA. After charge reduction, the polydisperse aerosol is subsequently introduced into the nDMA. There, the particles are separated according to their EMD by application of a tuneable negative voltage to the centre electrode of the nDMA in combination with an orthogonal laminar flow of sheath air. Only the positively charged particles are attracted towards the inner electrode, whereas the negatively charged particles are repelled and the neutral particles leave the nDMA unaffected. As a consequence, merely positively charged analytes with a certain EMD can exit the nDMA at a particular applied voltage. (Lpm: litres per minute; according to [166])

In this regard, the EMD of a particle (or Millikan diameter) is defined as the diameter of a sphere with the same electrophoretic mobility as the particle and is also related to its electrophoretic mobility Z_P [181, 145]. If a charged particle is introduced into an electrical field E, it undergoes a movement caused by electrical forces F_{el}. It quickly reaches its terminal velocity v, when the frictional forces F_{fr} equal the electrical ones **(8, 9)**. This velocity is proportional to the electric field E with the proportionality

constant Z_P **(10)**. From that equation **(11)** can be calculated, which defines its dependence on the particle charge and diameter. These derivations apply for spherical particles and include the Cunningham slip correction factor C_c, which corrects the frictional forces for larger particles: If the sizes of the particles approach the mean free path of the gas, the F_{fr} is smaller than given by the mere Stokes' law [161].

$$F_{el} = F_{fr} : neE = 3\pi\eta v D_{EMD} / C_c \quad \textbf{(8)}$$

$$v = neEC_c / 3\pi\eta D_{EMD} \quad \textbf{(9)}$$

$$v = Z_P E \quad \textbf{(10)}$$

$$Z_P = neC_c / 3\pi\eta D_{EMD} \quad \textbf{(11)}$$

ne: n elementary charges; η: gas viscosity; D_{EMD}: EMD

Ultimately, the EMD is influenced by the charge of the particle, the nDMA dimensions, the sheath flow rate, and the applied voltage [182]. As the charge (singly charged), the sheath flow, and the nDMA are usually kept constant throughout an experiment, the EMD is directly dependent on the applied electric field and analytes can thus be size-selectively separated by varying the voltage on the nDMA. Scanning through the whole voltage range, on the other hand, allows for a complete size distribution analysis of a sample.

3.2.3 Detection and sampling after size-selective separation

For detection the size-selected analytes are guided into a continuous-flow alcohol or water driven CPC (**Figure 3.10a**) [183, 184]. There, they are individually counted by an optical system measuring scattered light of a laser diode. The n-butanol-driven CPC can detect particles after enlarging them to 10-12 µm droplets by condensation of the alcohol (or water) onto the droplet. The gaseous sample is saturated with the alcohol vapour by passing over heated liquid butanol. By cooling the mixture in a condenser, the analytes act as condensation nuclei for the butanol vapour. A controlled supersaturation below the critical saturation ratio is established in the CPC, in order to prevent self-nucleation, in which butanol condenses onto butanol clusters formed in air. This also limits the detection of particles below 3 nm, as they are not enlarged by condensation and thus not detected [185].

Next to mere detection, particles can be collected onto various substrates after gas-phase separation with an electrostatic nanoparticle sampler (ENAS). This allows for subsequent analysis of the analytes with e.g. microscopic measurements (transmission electron microscopy, TEM; atomic-force microscopy, AFM) or biological tests like immunological assays [186-188]. The ENAS consists of an electrically grounded sampling chamber that contains an electrode in its bottom centre (**Figure 3.10b**).

By application of a negative voltage up to 10 kV to this electrode, positively charged particles coming from the nDMA are attracted. Consequently, they are sampled onto a substrate (e.g. TEM grid, freshly cleaved mica platelet, or nitrocellulose membrane) mounted on top of the electrode with adhesive tape. The deposition rate is affected by the adjustable flow rate, with which the analytes enter the sampling chamber, by the applied voltage, as well as by the particle concentration and charge.

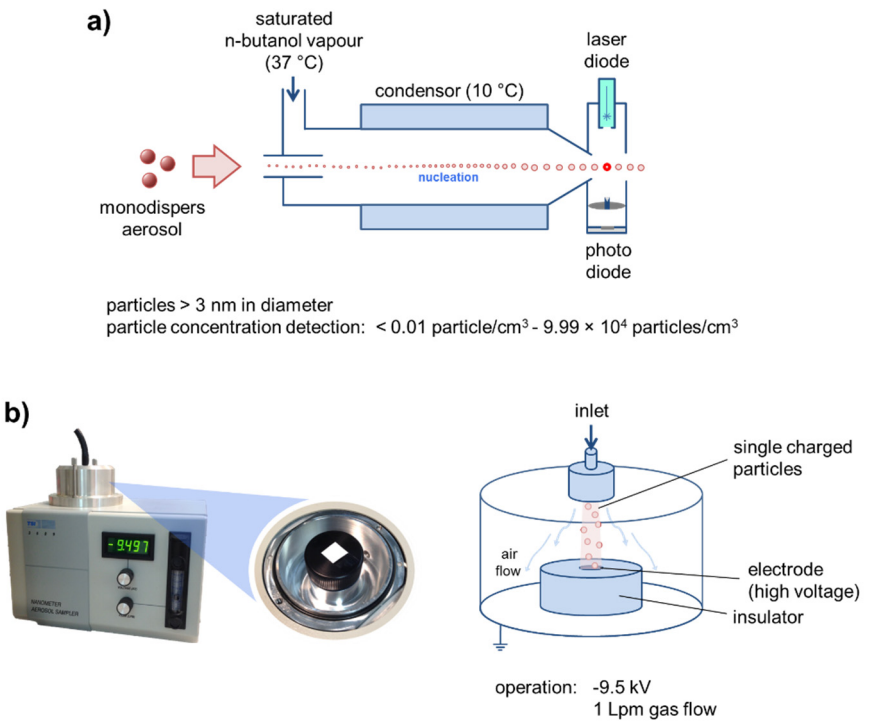


Figure 3.10 – Exemplary detection and sampling. For detection the obtained monodisperse aerosol is directed through a CPC, in which supersaturated water or alcohol (e.g. n-butanol) vapour condenses onto the particles (**a**). Following nucleation, single particles can be counted by laser light scattering yielding e.g. a number-based particle concentration. Alternatively, particles can be size-selectively collected with an ENAS after gas-phase separation for consecutive investigations like microscopic measurements or immunological assays (**b**). (according to [166, 186])

Aim of the thesis

The present thesis focussed on establishing new sensitive and specific methods of analysis for glycoproteins and their interactions with lectins. Therefore, the glycoproteins were investigated either directly or after incubation with lectins by electrophoresis in liquid and in gas-phase.

A set of well-known glycoproteins with varying degrees and types of glycosylation (transferrin, antitrypsin, and acid glycoprotein) was employed to assess the applicability of MCGE for the analysis of glycoproteins in liquid-phase (**Publication I**). Parameters like precision and repeatability of sizing (MW determination) and quantitation (both on- and off-chip), limit of detection and quantitation, as well as linear dynamic range were evaluated and compared to traditional SDS-PAGE results.

For the targeted analysis of glycoproteins in complex biological samples, MCGE was combined with a specific affinity enrichment using lectins covalently linked to magnetic beads (**Publication II**). The lectin coupling, the glycoprotein enrichment, and the successive analysis with MCGE were optimized with the same set of model glycoproteins and various model lectins (*Sambucus nigra* agglutinin, concanavalin A, and wheat germ agglutinin). Binding and elution conditions were chosen with regard to the lowest unspecific interactions. Several biological samples (human serum, human serum depleted of the twelve most abundant proteins, and a mycelia extract of fungus *Trichoderma atroviride*) were analysed with the developed strategy in order to proof the applicability of the method concerning specificity and selectivity. Results were compared and confirmed by SDS-PAGE followed by tryptic *in-gel* digestion and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS).

The chosen lectin-glycoprotein systems were further investigated by nES GEMMA in gas-phase (**Publication III**). Each glycoprotein and lectin was analysed individually for size-determination and results compared to MALDI-TOF-MS measurements. As nES GEMMA allows fragile non-covalent interaction analysis, also incubations of each lectin with each glycoprotein were studied. For comparison, the experiments were additionally conducted in liquid-phase performing CE-on-a-chip measurements.

Confirming the formation of the non-covalent lectin-glycoprotein complex in liquid-phase and its maintenance throughout gas-phase separation, the complex was sampled with nES GEMMA onto a nitrocellulose membrane. A successive dot-blot immunological assay was optimized with a set of two appropriate antibodies recognizing only the glycoprotein of the complex.

Publications

Electrophoresis (**2015**), *36,* 1754-1758

Challenges of glycoprotein analysis by microchip capillary gel electrophoresis

Nicole Engel[1], Victor U. Weiss[1], Christian Wenz[2], Andreas Rüfer[2], Martin Kratzmeier[2], Susanne Glück[2], Martina Marchetti-Deschmann[1], Günter Allmaier[1]

[1]Institute of Chemical Technologies and Analytics, Vienna University of Technology, Vienna, Austria
[2]Agilent Technologies, Waldbronn, Germany

Nicole Engel[1]
Victor U. Weiss[1]
Christian Wenz[2]
Andreas Rüfer[2]
Martin Kratzmeier[2]
Susanne Glück[2]
Martina Marchetti-Deschmann[1]
Günter Allmaier[1]

[1]Institute of Chemical Technologies and Analytics, Vienna University of Technology, Vienna, Austria
[2]Agilent Technologies, Waldbronn, Germany

Received October 27, 2014
Revised March 24, 2015
Accepted April 9, 2015

Short Communication

Challenges of glycoprotein analysis by microchip capillary gel electrophoresis

Glycosylations severely influence a protein's biological and physicochemical properties. Five exemplary proteins with varying glycan moieties were chosen to establish molecular weight (MW) determination (sizing), quantitation, and sensitivity of detection for microchip capillary gel electrophoresis (MCGE). Although sizing showed increasing deviations from literature values (SDS-PAGE or MALDI-MS) with a concomitant higher degree of analyte glycosylation, the reproducibility of MW determination and accuracy of quantitation with high sensitivity and reliability were demonstrated. Additionally, speed of analysis together with the low level of analyte consumption render MCGE attractive as an alternative to conventional SDS-PAGE.

Keywords:
Chip electrophoresis / Glycoprotein / Lab-on-a-chip / Laser-induced fluorescence / SDS-PAGE
DOI 10.1002/elps.201400510

For decades, SDS-PAGE was the method of choice to separate and characterize protein samples according to their molecular weight (MW) [1, 2]. With the development of a chip-based microfluidic system for gel electrophoretic separations (MCGE, Agilent 2100 Bioanalyzer, Caliper LC90, and BioRad Experion) based on LIF detection, a rapid and easy-to-use method for sizing and quantitation of proteins with high reproducibility and sensitivity as alternative to the classical gel-based method was offered [3]. In contrast to SDS-PAGE a sample is analyzed in only one minute via MCGE with data acquisition in real-time. Comparisons of both methods have proven the applicability of MCGE and its comparability to SDS-PAGE concerning accurate sizing of plain proteins with the benefit of simultaneous quantitation and detection of impurities [4–6].

The 2100 Bioanalyzer MCGE system offers two assays with different labeling strategies for proteins. In both cases, analytes are separated in a sieving matrix of an entangled linear polymer in a SDS containing BGE. For the Protein 230 (P230) Assay, a fluorescent dye binds noncovalently to SDS-protein complexes during chip electrophoresis [7, 8]. An BGE dilution step before LIF detection is included to increase sensitivity [3]. In contrast, the High Sensitivity Protein 250 (HSP-250) Assay employs a covalently linked fluorescence label introduced to the sample prior MCGE, resulting in a sensitivity comparable to and sometimes even exceeding silver-staining [5]. Besides the extended linear dynamic range of the HSP-250, both assays exhibit similar analytical characteristics (reproducibility and accuracy of quantitation and sizing) for unmodified standard proteins.

However, the analysis of considerable posttranslational modified proteins, especially glycoproteins, has recently gained enormous interest. Besides their high occurrence and diversity, glycan modifications have a major impact on proteins: next to influencing a protein's physicochemical properties and biological activity, glycosylations play an important role in recognition events, protein trafficking, immunology, and cancer biology [9, 10].

This Short Communication addresses the properties of MCGE for detailed glycoprotein analysis. Five proteins — bovine alpha-1-acid glycoprotein (AGP), human alpha-1-antitrypsin (A1AT), recombinant human erythropoietin beta (EPO), hen egg white ovalbumin (OA), and human serum transferrin (TF) — with different degrees and patterns of glycosylation were separated with the P230 and the HSP-250 Assay and results were compared to SDS-PAGE. Peak pattern and shape, reproducibility and accuracy of sizing and quantitation, as well as sensitivity of detection were of interest.

TF (≥ 98%), OA (≥ 98%), AGP (99%), and A1AT (salt free, lyophilized powder) were purchased from Sigma-Aldrich (Taufkirchen, Germany) and EPO (NeoRecormon®, 500 IU EPO-β/0.3 mL) from Roche. Proteins were analyzed on the Agilent 2100 Bioanalyzer system (Waldbronn, Germany) either with the P230 or HSP-250 Assay according to the manufacturer's protocols. Briefly, glycoproteins were dissolved in water (700 μg/mL stock solution) and incubated (95°C, 5 min) with sample buffer containing DTT (Sigma-Aldrich). For the P230 Assay samples were 15-fold diluted with water before analysis. For the

Correspondence: Professor Günter Allmaier, Institute of Chemical Technologies and Analytics, Vienna University of Technology, Getreidemarkt 9/164-IAC, 1060 Vienna, Austria
E-mail: guenter.allmaier@tuwien.ac.at
Fax: +43-1-58801-15199

Abbreviations: AGP, alpha-1-acid glycoprotein; **A1AT**, alpha-1-antitrypsin; **EPO**, recombinant erythropoietin; **MCGE**, microchip capillary gel electrophoresis; **MW**, molecular weight; **OA**, ovalbumin; **TF**, transferrin; **wb**, baseline width

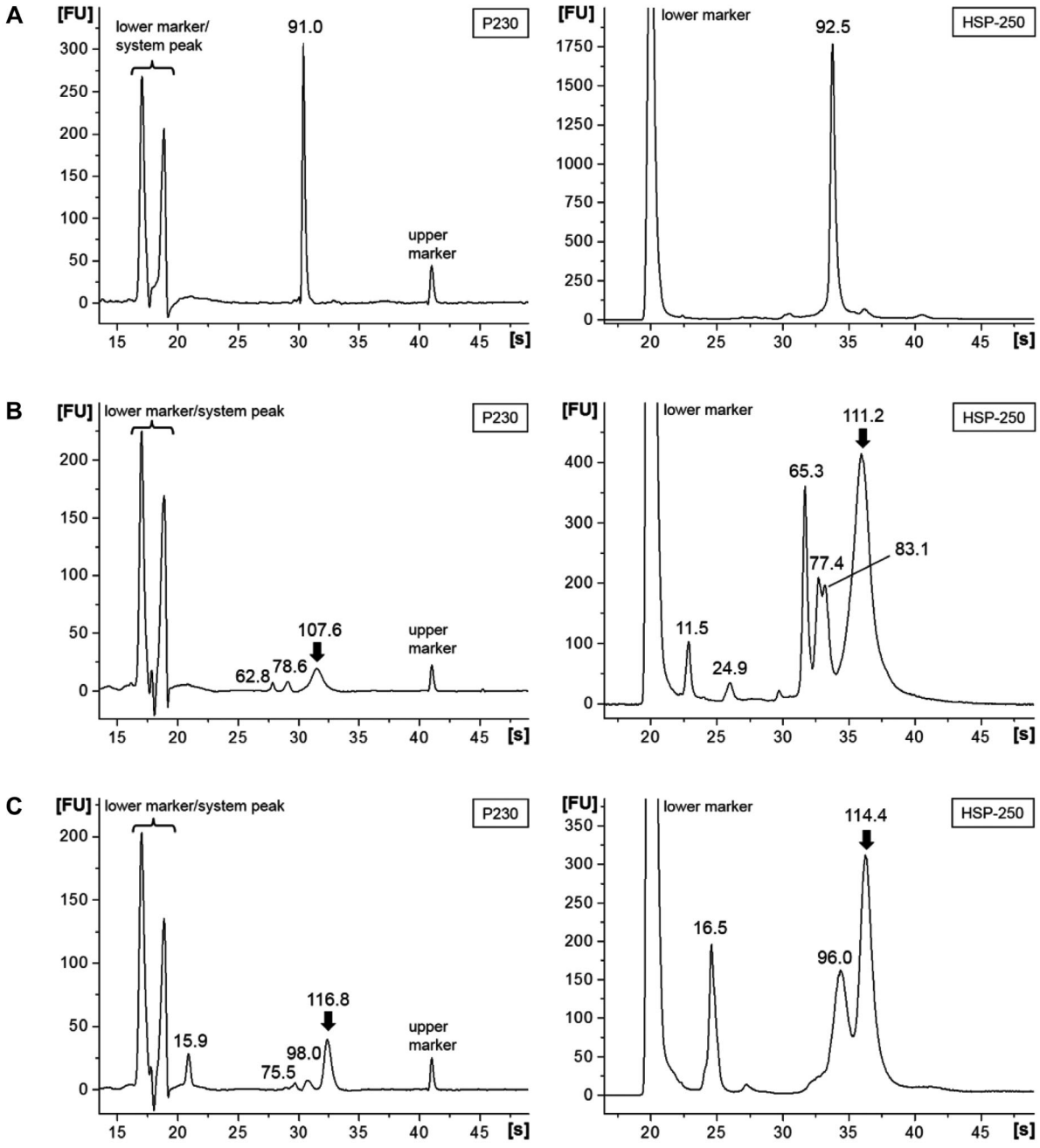


Figure 1. MCGE analysis of model glycoproteins (550 μg/mL each) comparing the P230 and the HSP-250 Assay. TF (A), A1AT (B), and AGP (C) were separated on the 2100 Bioanalyzer with the HSP-250 Assay exhibiting a higher sensitivity (50, 100, and 25 times lower LODs for the most abundant signals marked with an arrow).

HSP-250 Assay samples were diluted 1:200 with water and, prior to denaturation, were fluorescence labeled according to protocol (EPO was labeled overnight). For SDS-PAGE separations samples in LDS sample buffer and 50 mM DTT were run on NuPAGE 4–12% Bis-Tris gels in MES SDS running buffer (all from Invitrogen, Darmstadt, Germany) at 120 V (const.) and 60 mA (max.). Protein bands were visualized by MS-compatible silver staining [11]. MS experiments were performed on a MALDI-TOF AXIMA TOF^2 instrument (Shimadzu Kratos Analytical, Manchester, UK) in linear positive ion mode at 20 keV acceleration using stainless steel sample plates (matrix: 12 mg/mL sinapic acid in 0.1% TFA/acetonitrile (1:1, v/v), dried-droplet technique).

OA and EPO are glycoprotein examples for extremely low (4%) and high (39%) glycan contents. OA (one N-glycosylation site, short antennas) showed similar behavior to unmodified

Table 1. Analysis of OA, TF, A1AT, AGP, and EPO by MALDI-MS, SDS-PAGE, and P230/HSP-250 Assays

Glycoprotein	Approx. N-glycosylation (w/w %)[a)]	MALDI-MS MW_{lit} (kDa)	MALDI-MS MW_{exp} (kDa)[b)]	SDS-PAGE MW_{exp} (kDa)[b)]	P230 MW_{exp} (kDa)[b)]	HSP-250 MW_{exp} (kDa)[b)]	P230/HSP-250 LOD (μg/mL)	P230/HSP-250 CV%[c)]	P230/HSP-250 R^2 [d)]
OA	4	44.5	40.1, **44.4**	40, **45**, 79	42.7 ± 0.5 **44.7** ± 0.5 85.6 ± 1.9	**43.8** ± 0.5 87.8 ± 1.9	25/1	9/4%	0.994/0.999
TF	6	80	**79.2**	**80**	**89.3** ± 1.3	**90.4** ± 1.4	25/0.5	9/10%	0.998/0.995
A1AT	13	51	**35.2**, 50.9, 57.3, 66.6	49, **61**, **68**	63.4 ± 1.0 78.8 ± 1.1 **108.1** ± 1.4	64.6 ± 0.8 77.2 ± 1.0 87.0 ± 1.0 **111.9** ± 1.1	50/0.5	8/10%	0.998/0.993
AGP	37	33.8	**13.6**, **31.9**, 45.5, 59.4	15, 43, **47**, 54	15.6 ± 0.2 84.7 ± 0.7 98.5 ± 1.2 **116.9** ± 1.2	15.8 ± 0.4 95.2 ± 1.3 **113.4** ± 1.3	25/1	8/11%	0.992/0.994
EPO	39	30	**30.4**	**40**	**89.8** ± 1.3	**80.8** ± 2.9	35/1	8/8%	0.990/0.997

a) Values according to references.
b) Bold numbers are referring to dominating glycoprotein species.
c) Average deviation of quantitation concerning the evaluated signal.
d) Correlation coefficient (R^2) of the linear regression plotting analyte concentration against time corrected area.

proteins during MCGE, sharp peaks, and accurate MW determination. This demonstrates that a low glycosylation degree has no influence in MCGE. In contrast, EPO (three N-glycosylations, one O-glycosylation) exhibits a very broad peak with high deviations in sizing (>200%). In case of the HSP-250 Assay even labeling conditions had to be adapted because the bulky glycan chains hindered a successful derivatization. TF, A1AT, and AGP represent glycoproteins of average size and glycosylation degree. Therefore, the following short communication concentrates on these glycoproteins to assess the suitability of MCGE for glycoprotein analysis in industry (biopharmaceuticals) and research.

Figure 1 displays the glycoproteins' MCGE separations with the P230 and the HSP-250 Assay. TF (697 amino acids) exhibits only two N-glycosylation sites and thus an average glycosylation content of about 6 w/w % [12]. Electropherograms (Fig. 1A) show a sharp, single signal at 89.3 ± 1.3 kDa for the P230 and 90.4 ± 1.4 kDa for the HSP-250 Assay, as determined from over 15 intra- and interchip measurements. Both peaks feature a small peak width at baseline (w_B: 1.2 ± 0.0/2.1 ± 0.1 s, P230/HSP-250) comparable to that of nonmodified proteins. In contrast, the signals of A1AT and AGP for both assays appear much broader resulting from enhanced heterogeneities caused by higher glycan contents (Fig. 1B and C). A1AT (394 amino acids, three N-glycans) has an approximate glycan content of 13 w/w % [13]. Its separation with the P230 Assay leads to the detection of three baseline separated peaks with MWs of 63.4 ± 1.0, 78.8 ± 1.1, and 108.1 ± 1.4 kDa, whereas more peaks could be resolved with the HSP-250 Assay (11.6 ± 0.1, 24.8 ± 0.4, 64.6 ± 0.8, 77.2 ± 1.0, 87.0 ± 1.0, and 111.9 ± 1.1 kDa). In both cases the signal of the largest component was most abundant and used for further evaluations (w_B: 3.3 ± 0.2/8.0 ± 1.4 s, P230/HSP-250). The intensities of the remaining peaks varied in both assays as a consequence of the different labeling strategies. Exhibiting five N-glycosylation sites on a 184 amino acid backbone, AGP is a protein with a high glycan content of about 37 w/w % [14]. The P230 Assay revealed five signals with MWs of 15.6 ± 0.2, 75.3 ± 0.6, 84.7 ± 0.7, 98.5 ± 1.2, and 116.9 ± 1.2 kDa, whereas only three distinct peaks could be separated with the HSP-250 Assay (15.8 ± 0.4, 95.2 ± 1.3, and 113.4 ± 1.3 kDa). Again the most abundant peak (w_B: 1.9 ± 0.2 / 4.2 ± 0.3 s, P230/HSP-250) was employed for further data evaluation.

Comparing MCGE derived MWs of the glycoproteins to MS data, results for all three analytes exceeded literature values (Table 1). Moreover, these deviations clearly showed a correlation with the amount of glycosylation: TF with the lower glycan content exhibited the smallest deviation of 12%, A1AT already revealed a difference of over 100%, and the MW of AGP exceeded the MS value by as much as 200%. Small differences in MW determination between the two MCGE assays result from a different analyte migration due to covalently linked dye molecules for the HSP-250 Assay.

The problem of inaccurate sizing is common for planar gel electrophoresis employing detergents and is thus also observed for SDS-PAGE of glycoproteins [15]. The uniform protein covering with SDS is altered leading to a slower migration in the gel compared to unmodified proteins and, consequently, to apparent higher MWs, since usually unmodified proteins are used as MW markers. Figure 2 exemplarily displays the MCGE electropherogram of A1AT, as well as its corresponding gel-like image and SDS-PAGE separation. The slab gel electrophoresis revealed three distinct bands with average MWs of 49, 61, and 68 kDa showing intensity variations compared with the HSP-250 Assay due to different staining

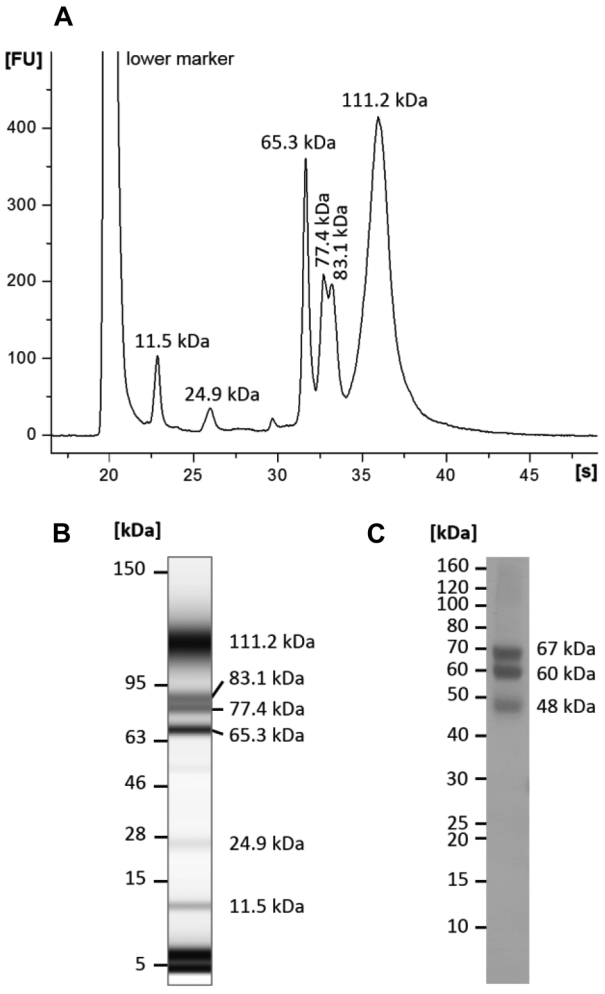


Figure 2. Comparison of the electropherogram (A) and gel-like image (B) of a HSP-250 MCGE analysis (550 μg/mL A1AT) with a SDS-PAGE separation of 250 μg of the same analyte (C).

techniques. It is of note that also SDS-PAGE results exceeded MS determined values (Table 1), however, to a lesser extent than MCGE.

Despite difficulties concerning sizing accuracy, the reproducibility of MW determination via MCGE was very high. Multiple sample triplicates of each analyte were measured to evaluate intra- and interchip variations. Values deviating less than 1.3% for intra- and 1.9% for interchip measurements, lie perfectly within specifications for unmodified proteins for both assays (3% sizing reproducibility for BSA), which is of great importance for quality control of recombinant glycoproteins.

To investigate the quantitation accuracy and reproducibility, triplets of different concentrations for each analyte were analyzed with the P230 and HSP-250 Assay, respectively, and plotted against time corrected areas to evaluate the correlation coefficients (R^2). All glycoproteins showed good linearity within this dynamic range with $R^2 > 0.99$ (Table 1). Furthermore, all analytes could be quantified within the specifications of both assays (quantitation reproducibility with CV% $<$ 20% for BSA, Table 1). Deviations were lower in separations with the P230 Assay (8 to 9%) and slightly increased (10 to 11%) in correlation to increasing glycosylation degrees in case of the HSP-250 Assay. In all cases, interchip were higher than intrachip variations, but always $<$ 16% for TF and $<$ 18% for A1AT and AGP.

Evaluating the sensitivities of both assays, the LOD and LOQ, defined as concentrations with a S/N exceeding 2 or 3, were determined. The HSP-250 Assay showed a higher sensitivity than the P230 Assay as for unmodified proteins. Values of 1.0/5.0 μg/mL (LOD/LOQ) for AGP were 25 to five times higher, values of 0.5/1.0 μg/mL (LOD/LOQ) for A1AT were 100 to 50 times higher, and values of 0.5 μg/mL (LOD and LOQ) for TF were 50 times higher. These numbers again reveal a correlation with the glycosylation degree for the HSP-250 Assay: decreasing sensitivity for higher modified proteins. The sensitivity of the P230 Assay appeared independent from the glycosylation degree.

In summary, analytical parameters of MCGE glycoprotein analyses showed a correlation with the degree of glycan modification. In particular, peak shape and sizing accuracy were influenced. Deviations to MS data are well known and caused by altered migration of modified proteins [16]. Nevertheless, the high reproducibility of sizing, as well as the accurate and reliable quantitation of analytes together with short analysis times as well as low instrument investment and sample consumption proves the applicability of MCGE for glycoprotein analysis. As a highly standardize, rapid, and sensitive method MCGE is especially useful for process monitoring and quality control in purity and integrity analyses of biopharmaceuticals and biosimilars.

Financial support was provided by an unsolicited Agilent Technologies - University Relations Program grant and an Austrian Science Foundation (FWF project TRP 29-N20) grant.

The authors have declared no conflict of interest.

References

[1] Laemmli, U. K., *Nature* 1970, *227*, 680–685.

[2] Shapiro, A. L., Viñuela, E., Maizel Jr, J. V., *Biochem. Biophys. Res. Commun.* 1967, *28*, 815–820.

[3] Bousse, L., Mouradian, S., Minalla, A., Yee, H., Williams, K., Dubrow, R., *Anal. Chem.* 2001, *73*, 1207–1212.

[4] Kuschel, M., Neumann, T., Barthmaier, P., Kratzmeier, M., *J. Biomol. Tech.* 2002, *13*, 172–178.

[5] Wenz, C., Marchetti-Deschmann, M., Herwig, E., Schröttner, E., Allmaier, G., Trojer, L., Vollmer, M., Rüfer, A., *Electrophoresis* 2010, *31*, 611–617.

[6] Mueller, R., Marchetti, M., Kratzmeier, M., Elgass, H., Kuschel, M., Zenker, A., Allmaier, G., *Anal. Bioanal. Chem.* 2007, *389*, 1859–1868.

[7] Jin, L. J., Giordano, B. C., Landers, J. P., *Anal. Chem.* 2001, *73*, 4994–4999.

[8] Steinberg, T. H., Jones, L. J., Haugland, R. P., Singer, V. L., *Anal. Biochem.* 1996, *239*, 223–237.

[9] Varki, A., *Glycobiology* 1993, *3*, 97–130.

[10] Dwek, R. A., *Chem. Rev.* 1996, *96*, 683–720.

[11] Shevchenko, A., Wilm, M., Vorm, O., Mann, M., *Anal. Chem.* 1996, *68*, 850–858.

[12] Gimenez, E., Benavente, F., Barbosa, J., Sanz-Nebot, V., *Rapid Commun. Mass. Spectrom.* 2007, *21*, 2555–2563.

[13] Sturiale, L., Barone, R., Palmigiano, A., Ndosimao, C. N., Briones, P., Adamowicz, M., Jaeken, J., Garozzo, D., *Proteomics* 2008, *8*, 3822–3832.

[14] Nakano, M., Kakehi, K., Tsai, M. H., Lee, Y. C., *Glycobiology* 2004, *14*, 431–441.

[15] Mechref, Y., Madera, M., Novotny, M. V., *Methods Mol. Biol.* 2008, *424*, 373–396.

[16] Westermeier, R. (Ed.), *Electrophoresis in Practice*, Wiley-VCH, Weinheim , Germany 1996, p. 120.

Analytical and Bioanalytical Chemistry (to be submitted **2017**)

Microchip capillary gel electrophoresis combined with lectin affinity enrichment employing magnetic beads for glycoprotein analysis

Nicole Y. Engel[1], Victor U. Weiss[1], Christian Wenz[2], Susanne Glück[2], Andreas Rüfer[2], Martin Kratzmeier[2], Martina Marchetti-Deschmann[1], Günter Allmaier[1]

[1]Institute of Chemical Technologies and Analytics, TU Wien, Vienna, Austria
[2]Agilent Technologies, Waldbronn, Germany

***Analytical Bioanalytical Chemistry,* to be submitted (2017)**

Microchip capillary gel electrophoresis combined with lectin affinity enrichment employing magnetic beads for glycoprotein analysis

Nicole Y. Engel[1], Victor U. Weiss[1], Christian Wenz[2], Susanne Glück[2], Andreas Rüfer[2], Martin Kratzmeier[2], Martina Marchetti-Deschmann[1], Günter Allmaier[1]

[1]Institute of Chemical Technologies and Analytics, TU Wien (Vienna University of Technology), Vienna, Austria

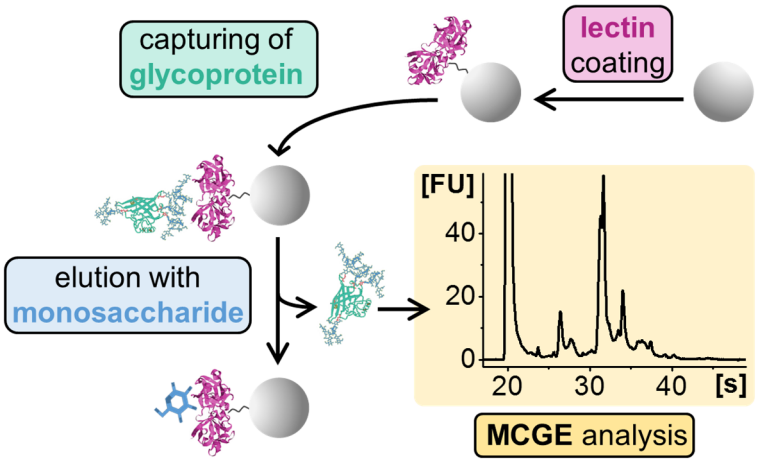
[2]Agilent Technologies, Waldbronn, Germany

Correspondence: Prof. Günter Allmaier, Institute of Chemical Technologies and Analytics, TU Wien, Getreidemarkt 9/164-IAC, 1060 Vienna, Austria

E-mail address: guenter.allmaier@tuwien.ac.at

Telephone: +43 - 1 - 58801 - 15160

Fax: +43 - 1 - 58801 - 15199



Graphical Abstract

Glycoproteins from biological samples were detected by microchip capillary gel electrophoresis after lectin affinity enrichment using magnetic beads and elution with respective competitive monosaccharides

Abstract

Due to the constant search for reliable methods to investigate glycoproteins in complex biological samples, an alternative approach combining affinity enrichment with rapid and sensitive analysis on-a-chip is presented. Glycoproteins were specifically captured by lectin-coated magnetic beads, eluted by competitive sugars, and investigated with microchip capillary gel electrophoresis (MCGE), i.e. CGE-on-a-chip. We compared our results to sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) data, which turned out to be in very good agreement. While SDS-PAGE offers the possibility of subsequent mass spectrometric analysis of captured analytes, MCGE scores with time savings, higher throughput, and lower sample consumption. Due to these advantages a lectin-based glycoprotein capture protocol was developed. Two different types of magnetic beads were tested and compared regarding lectin binding. The selectivity of our strategy was demonstrated with a set of model glycoproteins, as well as with human serum and serum depleted from high-abundance proteins. The specificity of the capturing method was investigated revealing an unspecific binding to a certain degree between each sample and the beads themselves, which has to be considered for any specific enrichment and data interpretation. In addition, two glycoproteins from *Trichoderma atroviride*, a fungus with mycoparasitistic activity and only barely studied glycoproteome, were enriched by means of a lectin and so identified for the first time.

Keywords: Affinity enrichment / Glycoprotein / Lab-on-a-chip / Lectin / Magnetic beads

Introduction

Glycosylations are regarded the most complex and, on the other hand, most common type of post-translational modifications with more than 50 % of all eukaryotic proteins being glycoproteins [1]. Glycoprotein analysis can be difficult due to their considerable macro- and microheterogeneity: Attached sugar moieties can range from single monosaccharides to complex linear or branched oligo- or even polysaccharides. Their huge structural diversity is also reflected in the variety of their functional purposes, which include their role in many molecular recognition events [2].

Carbohydrate binding proteins like lectins play nowadays an important role in the structural and functional elucidation of glycoproteins, as well as in the study of their binding affinities and interactions with other proteins [3, 4]. In contrast to antibodies, lectins are generally more stable, more affordable, better characterized, and address a broader spectrum of glycoproteins. On the other hand, they have lower binding specificities, which, however, can be of advantage for certain analytical strategies and application areas. One of their main fields of application is lectin affinity chromatography, during which they can be used for the specific and selective isolation of certain glycoproteins of interest [5, 6]. Therefore, lectins are immobilized on e.g. agarose- or silica-based media and filled into chromatographic columns of quite different dimensions. Typical lectins in this context are Concanavalin A (ConA), wheat germ agglutinin (WGA), and *Sambucus nigra* agglutinin (SNA). ConA specifically recognizes the trimannosidic core structure of an N-glycoprotein and other high-mannose structures [7], WGA binds to terminal N-acetylglucosamine and its β(1,4)-linked oligomers [8], and SNA has a high affinity for sialic acids (N-acetylneuraminic acid) α-glycosidically linked to galactose or N-acetylgalactosamine [9].

Today also other solid supports like magnetic beads are used especially for glycoprotein enrichment from crude biological samples without a chromatographic set-up [10]. They allow working directly from small sample amounts without elaborate sample preparation and simplify sample washing as well as enrichment. In order to enrich a great variety of glycoproteins, lectins having a broader specificity can be combined [11]. Therefore, a combination of e.g.

ConA, WGA, and SNA enables the enrichment of a wide range of different N-glycoproteins based on the lectins' specificities for common structural N-glycan moieties.

For investigation of isolated glycoproteins, the specific lectin affinity enrichment step has to be combined with an appropriate analytical technique. Very often gel electrophoresis, especially sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE), is chosen for a first insight into enrichment efficiency. Although being a generally established method and allowing subsequent analyte identification via mass spectrometry after separation, SDS-PAGE has some major draw-backs. Next to being very time-consuming, it lacks automation and therefore high throughput, requires certain technical skills, i.e. manual handling, and the uniform staining of glycoproteins after separation can be influenced by their modifications hampering quantification. Microchip capillary gel electrophoresis (MCGE), on the other hand, offers rapid (1 min per sample) and reliable size separations in addition to be performed in a high throughput fashion and quantitation in real-time [12-14]. As a microfluidic system it combines sample handling, separation of analytes, staining, laser-induced fluorescence (LIF) detection, and data analysis on one chip. Depending on the type of fluorescence labelling, available protein assays feature sensitivities comparable to silver-stained (covalent pre-chip labelling) or Coomassie-stained (dynamic on-chip labelling) gels. Both available MCGE protein assays have already been characterized in regard to glycoprotein sizing, limits of detection, and quantitation [15].

In this work, glycoprotein investigation by MCGE was successfully combined with a preceding specific lectin affinity enrichment using magnetic beads. Due to the speed of MCGE analysis, its low sample consumption and the ease of quantification of obtained results it was possible to study lectin affinity enrichment protocols in detail. For optimization of the presented approach two magnetic beads from different providers based on different basic chemical compostion (silica and polymer) were tested and compared, namely MagSi-S Tosyl 1.0 (MagSi) beads and tosylactivated Dynabeads MyOne (Dynabeads). Based on preliminary investigations detailed binding and elution conditions were evaluated finally for Dynabeads with a set of model lectins and glycoproteins of varying molecular weight and degree of glycosylation. For validation of the method and investigation of its selectivity, complex

biological samples (human blood serum and mycelia derived from *Trichoderma atroviride*) were used. All results were compared to traditional SDS-PAGE analysis. Additionally, enriched glycoproteins were identified by tryptic *in-gel* digestion and subsequent matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS).

Materials and Methods

Materials

Human serum transferrin (TF, ≥ 98 %), bovine acid glycoprotein (AGP, 99 %), human antitrypsin (A1AT, salt free, lyophilized powder), β-galactosidase from *E. coli* (β-Gal, lyophilized powder), bovine serum albumin (BSA, ≥ 96 %), as well as all other chemicals (purity of at least 99 %, each) were purchased from Sigma-Aldrich (St. Louis, MO, USA), if not stated otherwise. Sodium thiosulfate pentahydrate, ammonium sulfate, ethanol, acetonitrile (ACN), and acetic acid (all analytical grade) were obtained from Merck (Darmstadt, Germany). NuPAGE 4-12 % Bis-Tris gels, 4x lithium dodecyl sulfate (LDS) sample buffer (106 mM Tris HCl, 141 mM Tris Base, 2 % LDS, 10 % glycerol, 0.51 mM EDTA, 0.22 mM Serva Blue G250, 0.175 mM Phenol Red, pH 8.5), 20x MES SDS running buffer (50 mM MES, 50 mM Tris Base, 0.1 % SDS, 1 mM EDTA, pH 7.3), and BenchMark Protein Ladder were acquired from Life Technologies (Carlsbad, CA, USA). ConA, WGA, and SNA were obtained from Vector Laboratories (Burlingame, CA, USA). Dynabeads MyOne Tosylactivated were from Invitrogen Dynal (Oslo, Norway) and MagSi-S Tosyl 1.0 beads from MagnaMedics (Geleen, The Netherlands). The Agilent 2100 Bioanalyzer, the Protein 230 (P230) and the High Sensitivity Protein 250 (HSP-250) Kit were obtained from Agilent Technologies (Waldbronn, Germany). Sequencing grade trypsin from bovine was from Roche (Mannheim, Germany). All experiments were performed employing water of Millipore grade (18.2 MΩcm resistivity at 25°C) taken from a Simplicity system (Millipore, Molsheim, France).

Buffers

Lectins were dissolved in 0.1 M sodium borate and 0.1 mM $CaCl_2$ (pH 8.3) to 100 µM and stored at 4 °C until usage. Coupling buffer (0.1 M sodium borate, 1 M ammonium sulfate, pH 9.5) was freshly prepared. Blocking buffer (1 M Tris, 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$, pH 7.4) and binding buffer (20 mM Tris, 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$, pH 7.4) were prepared and stored at 4 °C for no longer than two weeks. pH was checked before usage.

Sample preparation

Human blood was taken from a healthy, voluntary donor with a sterile lancet and centrifuged at 14000 g, for 30 min. The serum was stored at -20 °C. According to the manufacturer's operating procedure, 10 µl of serum were incubated in a Pierce Top 12 Abundant Protein Depletion Spin Column (Thermo Fisher Scientific, Waltham, USA) for 1 h to deplete high-abundance proteins. The filtrate was concentrated and its buffer exchanged to binding buffer using 10 kDa Millipore Microcon centrifugal filters (Merck Millipore, Darmstadt, Germany) with an Ultracel regenerated cellulose membrane.

In the case of *T. atroviride*, frozen mycelia was equilibrated to ambient temperature before analysis. 100 mg of wet cell mycelia were suspended in 1.6 ml binding buffer and lysed by sonication (intensity 60 %, 2× 20 s and 1× 30 s, 1 min waiting intervals each, at 4 °C) with a Branson Sonifier 250 (Branson Ultrasonics, Danbury, CT, USA). Lysed cells were centrifuged (14000 g, 4 °C, 20 min), the supernatant was collected and concentrated by means of 3 kDa Millipore Microcon centrifugal filters.

Protein concentrations of all samples were determined using Bradford Assay and BSA for calibration.

Glycoprotein enrichment

Magnetic beads were handled as suggested by the manufacturer. Briefly: 5.33 µM lectin were covalently linked to 20 µg/µl tosylactivated magnetic beads in coupling buffer at 37 °C and 600 rpm for 20 h. Free tosyl groups were inactivated by overnight incubation in blocking buffer

at 37 °C and 600 rpm. The prepared beads were washed two times with a fivefold volume of binding buffer and stored in the same buffer at 4 °C. Prepared beads were found to be stable for at least two weeks and were used within this time period.

Glycoprotein samples (0.5 µg/µl standard proteins, 4.7 µg/µl serum, 2.0 µg/µl depleted serum, 4.7 µg/µl *T. atroviride*) were incubated with 20 µg/µl lectin beads in binding buffer at room temperature and 600 rpm for 60 min to specifically capture the glycoproteins. The beads with captured glycoproteins were washed three times with fivefold volume of binding buffer. Glycoproteins were eluted from the beads with the respective competitive mono- and disaccharide of the lectin at room temperature for 15 min and 600 rpm. Methyl α-D glucopyranoside and methyl α-D mannopyranoside, 200 mM each, were used for ConA, 500 mM N-acetyl-D-glucosamine for WGA, and 500 mM lactose for SNA. Elution of remaining non-covalently bound analytes was achieved by incubating the beads for 5 min at 95 °C either with 5 µl 4x LDS sample buffer containing 50 mM DTT for SDS-PAGE or with 6 µl P230/HSP-250 sample buffer (diluted according to manufacturer's protocol in water) containing 11.7 mM DTT for MCGE.

MCGE

The chip-based glycoprotein analysis was performed on the Agilent 2100 Bioanalyzer system either with the P230 or the HSP-250 assay according to the manufacturer's instructions. Briefly, samples were fluorescently labelled for the HSP-250 assay prior to denaturation. 10 µL of sample solution were mixed with 1 µL HSP-250 labelling dye and incubated on ice for 30 min. The labelling reaction was stopped by adding 1 µL of ethanolamine followed by incubation on ice for additional 10 min. No labelling was necessary for samples analysed by the P230 assay. Samples were diluted only with water prior to denaturation. For this, 4 µL sample were incubated (95 °C, 5 min) with 2 µL sample buffer containing DTT. HSP-250 assay samples were directly applied to the chip, whereas P230 assay samples were diluted with 84 µL water before application.

SDS-PAGE

SDS-PAGE analyses were carried out on NuPAGE 4-12 % Bis-Tris gels using MES SDS running buffer at 120 V (const.) and 60 mA (max.). BenchMark Protein Ladder was applied for molecular weight determination. Protein bands were visualized by silver staining suited for further MS analyses [16]. Briefly, gels were washed with an aqueous solution containing 50 % ethanol and 5 % acetic acid for 20 min, with 50 % ethanol for 10 min, and three times with water for 20 min each. Afterwards, gels were incubated in 0.2 % sodium thiosulfate pentahydrate for 1 min and washed two times with water for 1 min. Subsequently, gels were incubated in 1 % silver nitrate at 4 °C for 20 min and washed twice with water for 1 min. Gels were further incubated in 2 % sodium carbonate containing 0.04 % formaldehyde until protein bands were visible. Staining was stopped with 5 % acetic acid. Gels were stored in 1 % acetic acid at 4 °C for further investigations.

Tryptic *in-gel* digestion and MALDI mass spectrometric identification

After SDS-PAGE glycoproteins were identified by MALDI-RTOF-MS and -MS/MS analyses after destaining the gel and tryptic *in-gel* digestion. Briefly, excised gel bands were destained in 100 mM $Na_2S_2O_3$ / 30 mM $K_4Fe(CN)_6 \cdot 3H_2O$ (1:1, v/v) [17]. Gel pieces were treated with ACN, rehydrated (100 mM NH_4HCO_3), reduced (10 mM DTT in 100 mM NH_4HCO_3, 56 °C, 45 min), alkylated (50 mM iodoacetamide in 100 mM NH_4HCO_3, 24 °C, 30 min), and finally dried in a vacuum centrifuge. After rehydration in approx. 10 µL 50 mM NH_4HCO_3 (pH 8.5) containing 5 % ACN and 100 ng trypsin and incubation at 37 °C overnight, peptides were extracted with 50 mM NH_4HCO_3/ACN (1:1, v/v) and 0.1 % formic acid / ACN (1:1, v/v). All extracts of selected lanes were dried in a vacuum centrifuge. After reconstitution in 0.1 % aqueous TFA, peptides were desalted using C_{18} ZipTip® pipette tips (Merck Millipore) and eluted with 2.7 µg/µL α-cyano-4-hydroxycinnamic acid (MALDI-MS matrix) prepared in aqueous ACN (50 %) containing 0.1 % TFA. Peptide mass fingerprinting (PMF) and MS/MS analyses were carried out on an UltrafleXtreme MALDI-linear TOF/RTOF instrument using an AnchorChip target (both Bruker Daltonik, Bremen, Germany).

For all data related to enzymatic digestion, autolytic tryptic products, keratin, and blank artefacts were assigned and removed before database search (SWISSPROT with taxonomy *Homo sapiens* for human samples and in all entries of NCBInr for *Trichoderma* samples, September - December 2013) using Mascot [18] with the following parameters: monoisotopic mass values, peptide mass tolerance of ± 0.15 and 0.3 Da for PMF and MS/MS experiments, respectively, 1 missed cleavage, carbamidomethylation as fixed modification, and methionine oxidations and methylation of the protein N-terminus as variable modifications.

Results and discussion

A general scheme of the specific affinity enrichment of glycoproteins combined with MCGE analysis is presented in **Electronic Supplementary Material (ESM) Fig. S1**. In a first step the individual lectins ConA, WGA, and SNA were covalently coupled to tosylactivated magnetic beads. These lectin-coated beads were incubated with a sample containing glycoproteins, which had been fluorescently labelled for subsequent chip analysis. For a specific elution of the enriched glycoproteins the respective competitive mono- or disaccharides to the lectins were applied and the eluted glycoproteins were analysed with MCGE or, for comparison, with SDS-PAGE. A second unspecific elution step was added by incubation of the beads with denaturing buffer containing LDS or SDS to assess the degree of specific lectin and unspecific protein binding, respectively.

Selection of magnetic beads

Two different magnetic beads with nominal diameters of 1 µm according to manufacturers' specifications were tested and compared. However, scanning electron microscopy experiments of the two bead materials revealed a very broad size distribution of the MagSi beads ranging from about 100 nm up to 5 µm in contrast to the rather size-uniform Dynabeads (**Fig. 1a**). This circumstance can influence effectiveness of enrichment and reproducibility of analyses.

Coupling conditions and concentrations were optimized for both beads and binding efficiencies examined using the lectin ConA. Therefore, after binding ConA to the beads the supernatant (S) solution and the unspecific elution (E) fraction of the beads were analysed with SDS-PAGE and MCGE (**Fig. 1b** and **c**). Remaining ConA in the supernatant was used as an indicator for a sufficient surface coverage of the bead, whereas the unspecific elution of ConA was used for assessing the amount of non-covalently bound ConA. In the case of the Dynabeads, application of 5.33 µM ConA to 20 µg/µl beads showed a slight SDS-PAGE band in the supernatant after incubation, pointing to a sufficient lectin concentration to saturate all binding sites (**Fig. 1b**). Biologically active ConA is a non-covalent multimer (tetramer at physiological pH) [19]. Therefore, a protein band at the apparent molecular mass of the monomer (26 kDa) was expected to be characteristic for the unspecific elution of the beads. **Fig. 1b** shows no band in the supernatant and only a slight band in the unspecific elution fraction of the MagSi beads when applying the same amount of lectin. This indicates that ConA monomers were almost completely linked to the bead surface without the formation of non-covalently, biologically active multimers. Similar results were observed after doubling the lectin concentration while keeping the bead concentration constant: no remaining ConA in the supernatant, but a weak band in SDS-PAGE for the unspecific elution indicating the formation of biologically active lectin multimers. Only after application of a tenfold amount of ConA a distinct ConA band was visible in SDS-PAGE also for the supernatant indicating saturation of the magnetic bead surface. Quantitative analysis introduced by MCGE analysis allowed the determination of ConA in the supernatant and the unspecific elution fraction after generating a calibration function for different ConA concentrations (**Fig. 1c**). 45 % non-covalently bound ConA with regard to the initially applied amount was eluted in the case of Dynabeads and only 5 % in the case of MagSi beads. From these results it was concluded that Dynabeads need much less lectin while forming a higher number of biologically relevant multimers. Furthermore, first tests for the functionality of the immobilized lectins showed that MagSi beads exhibit a higher degree of unspecific protein binding (data not shown).

Based on these findings and with respect to the previously mentioned size heterogeneity of the MagSi beads further optimization was restricted to the use of Dynabeads.

Optimization of glycoprotein enrichment

For the optimization of the affinity enrichment a set of model glycoproteins (Tf, A1AT, and AGP) with varying degrees and types of glycosylations and the non-glycosylated protein β-Gal as negative control were applied. Tf was the largest of the applied glycoproteins in terms of molecular mass but exhibited the lowest glycosylation content (one O-glycan, two N-glycans, low degree of sialylation) [20]. The smaller A1AT featured a higher glycosylation degree (one O-glycan, three N-glycans, higher number of sialylation) [21]. AGP, the smallest glycoprotein, had the highest glycan content (five N-glycans) and most sialic acids attached [22]. In all cases, the binding buffer contained $CaCl_2$ and $MnCl_2$, as these cations are required by some lectins for a successful carbohydrate recognition [19]. Next to that, the buffer needed to be compatible with subsequent MCGE analyses concerning pH and concentrations of constituents. Different time spans (15 min up to 2 h) and temperatures (4 °C up to 29 °C) during binding and elution steps were tested together with the addition of detergents (0.1 % up to 1 % Tween 20, Tween 80, Thesit, Triton X-100) to buffers, addition of extra blocking reagents (serum albumin or competitive monosaccharides) to the blocking buffer to reduce unspecific analyte binding to beads, and variations in elution reagents (acetic acid, repeated sugar elution, SDS). Least unspecific binding and elution of β-Gal was found for 1 h binding and 15 min elution steps at room temperature with no addition of detergents. Furthermore, the elution using competitive monosaccharides proved to be most specific, selectively eluting the glycoproteins but not the non-glycosylated β-Gal (**ESM Fig. S2**).

However, SDS-PAGE analysis demonstrated that glycoprotein elution was not complete. An increase of sugar concentration during elution did not effectively influence these results (data not shown). Besides glycoproteins still bound to lectin beads after specific elution also unspecifically bound β-Gal and non-covalently bound ConA was detected in the unspecific elution fraction. In favour of the higher specificity, the incomplete glycoprotein elution by using competitive sugars was accepted.

Combination of affinity enrichment and MCGE analysis

Combining the optimized affinity enrichment with MCGE considerably reduced the required time for analysis, detection, and data evaluation (MCGE: 0.5 h, SDS-PAGE: > 3.5 h), as well as the applied sample amount by at least 50 %. Due to enrichment buffer compatibility with MCGE buffer requirements, no additional buffer exchanges were necessary before CGE-on-a-chip analysis. Therefore, all samples could be directly analysed after specific elution. For LIF detection samples were fluorescently labelled with a covalent HSP-250 fluorescent dye using a minimal labelling strategy. By labelling only the sample before the affinity enrichment, no lectin signal interfered with signals from eluted glycoprotein.

ESM Fig. S3 and S4 display the electropherograms for experiments including Tf, β-Gal, A1AT, and AGP employing ConA beads. MCGE results for samples of the initially applied glycoprotein concentrations, of the supernatant after incubation with the lectin beads, and of the specific elution fraction for the individual enrichments are shown. The direct comparison of the Tf and β-Gal enrichment (**ESM Fig. S3**) confirmed results from SDS-PAGE analyses. About 65 % of Tf were specifically bound to the ConA beads, but also β-Gal showed a certain unspecific binding towards ConA. However, almost only Tf was eluted during the specific elution step (time corrected areas, 50 : 1 = Tf : β-Gal).

In addition, for A1AT and AGP the enrichment of different sample constituents could be found. The fact that more than one signal was observed for the respective glycoproteins can be explained most likely by glycoforms, but can also result from sample impurities (**ESM Fig. S4**). Especially in the case of A1AT, the supernatant as well as the elution fraction reveal distinct changes in the relative signal intensities. From that, a stronger binding of one constituent (77.6 kDa) to ConA could be inferred. The same results were found with SDS-PAGE analysis (**ESM Fig. S2**).

Next to the individual affinity enrichments also a mixture of the glycoproteins with and without the negative control were incubated with ConA beats and analysed with MCGE and SDS-PAGE (**Fig. 2**). All glycoproteins could be identified before and after enrichment in the electropherograms, yet, with varying intensities depending on the presence of β-Gal. This became most obvious with Tf and AGP, which both showed lower signal intensities in the

specific elution fraction, when β-Gal was added. It was found that the Tf signal was reduced by approx. 45 %. For AGP an estimation was not possible because of β-Gal co-migration, however the reduction in signal intensity is visible. β-Gal, on the other hand, was only detectable at much lower concentration levels in the specific elution fraction, which was also confirmed by SDS-PAGE. β-Gal binding seemed to be increased in the presence of other glycoproteins compared to the individual incubation of β-Gal with ConA beads. This raised the question whether β-Gal really interacted with ConA or rather binds to other glycoproteins ("β-Gal piggy-back on glycoproteins"). Such a binding also influences the binding and elution of target proteins and thus explains lower signal intensities. Nevertheless, the eluted amount of unspecifically bound negative control was small compared to the specifically bound glycoproteins as indicated by the MCGE signal intensities and SDS-PAGE band abundance, respectively.

Selective enrichment of glycoproteins from biological samples

For the selective enrichment of glycoproteins from more complex biological samples, WGA and SNA beads were prepared together with ConA beads. All selected lectins have particular specificities to different carbohydrate moieties, and combining them allows to target a broad range of glycoforms. An equimolar mixture of the separately prepared lectin beads was used to target a high number of different glycoforms in human serum, a biological sample already well-studied in respect to glycoprotein enrichment [23]. Therefore, freshly taken blood was centrifuged, the supernatant fluorescently labelled and incubated with the mixture of lectin beads.

Fig. 3 displays the electropherograms of serum before affinity enrichment and of the specific elution fraction using a mixture of complementary sugars for each lectin. The protein profile before enrichment was dominated by a peak at 64 kDa, the most abundant protein in human serum, albumin (HSA). In contrast, the elution fraction was characterized by a completely changed peak pattern: several new components could be identified and the distinct HSA signal has strongly decreased. Unspecific interactions of HSA as a carrier protein with other components of serum, also with glycoproteins and lectins, was expected to interfere with

sample enrichment, and indeed a complete depletion of the HSA peak was not achieved. However, the unspecific binding and elution of HSA could be decreased by varying the amount of applied serum and lectin beads. (**ESM Fig. S5**). The lowest HSA signal relative to other constituents was achieved using 8 µl of labelled serum (equivalent to 2 µl pure serum) and 6 µl of lectin beads (ConA:SNA:WGA = 1:1:1, 2 µl each) diluted in binding buffer to a final volume of 30 µl. In that ratio 4 µl of labelled serum equals 70 µg of protein and 1 µl of lectin beads corresponds to 100 µg beads, giving a final ratio of 140 µg protein and 600 µg beads. Keeping the ratio constant, but changing the concentrations of the components also influenced the peak pattern. Reducing the serum and bead concentration by 25 % (4 µl labelled serum, 3 µL lectin beads, total volume 20 µl) resulted in very low over-all signal intensities, yet also a low HSA signal in regard to the other peaks. In contrast, increasing the concentrations by 12.5 % (12 µl labelled serum, 9 µL lectin beads, total volume 40 µl) showed a relatively increased HSA peak and were thus avoided (e.g. time corrected areas, 1 : 2.9 = peak at 60.5 kDa : HSA, **ESM Fig. S5a**). Using lower concentrations for enrichments at the same time reduced sample consumption. A good compromise was reached using 2 µl of all (the three different lectin beads) components (in a total volume of 30 µl with the serum) obtaining high signal intensities and a comparably weak HSA signal (e.g. time corrected areas, 1 : 1.4 = peak at 60.5 kDa : HSA, **Fig. 3**). Changing the serum/lectin ratio by increasing the amount of beads resulted in a lower specificity of the enrichment with high HSA peaks (**ESM Fig. S5b**). Therefore, a mixture of 8 µl labelled serum and 6 µl lectin beads in a total volume of 30 µl was used for further experiments.

The analysis of enriched glycoproteins from serum with SDS-PAGE and subsequent tryptic *in-gel* digestion and MALDI-RTOF-MS for protein identification further demonstrated the selectivity of the method (**Fig. 4a**). For a better comparability the ratios and amounts of used serum and lectin beads were kept the same as evaluated and mentioned before. Next to HSA, which was expected to be detected from previous MCGE experiments, only glycoproteins could be identified in the specific elution fractions. One has to be aware that the identified, non-glycosylated Ig kappa chain was enriched as part of the intact glycosylated antibody, but separated during SDS-PAGE because of reducing conditions.

Further evaluating the selectivity of the approach, the enrichment was applied to serum depleted from the twelve most abundant serum proteins. Therefore, the serum was treated with Top 12 spin columns before incubation with the mixture of lectin beads. **Fig. 4b** clearly shows the reduction of HSA at 69 kDa and of IgG at 56 and 27 kDa on the SDS-PAGE. Again only glycoproteins could be identified from the depleted serum.

As further additionally application, the cellulose degrading fungus *T. atroviride* was investigated with the established strategy. The ability of filamentous fungi to secrete large amounts of glycoproteins rose the interest to use the established lectin based enrichment approach. Although there is only little information about the nature of the glycoproteins in those fungi the predominant forms were found to be oligomannose N-glycans and O-glycans [24]. In contrast, more complex glycan structures as present in mammalians were not detected so far. Next to that, the presence of glucose, galactose, and N-acetylglucosamine has been reported on the glycans. Using again a mixture of the three lectin beads, especially ConA was expected to specifically bind to the high-mannose structures of the *T. atroviride* sample. WGA should show interactions with potentially present terminal N-acetylglucosamines. As mainly high-mannose-type glycoproteins were detected in filamentous fungi, interactions with SNA were not particularly expected. However, it was still included in the experiment in order to target a broader range of N-glycans.

The enrichment resulted in the identification of two proteins, glyceraldehyde-3-phosphate dehydrogenase and β-galactosidase. The latter is known to have several putative N-glycosylation sites in *T. reesei* [25]. Despite of the higher number of proteins being captured by the lectin beads, as shown in the unspecific elution fraction (**Fig. 4c**), only two were significantly eluted in the specific elution fraction. One explanation is a possible qualitative difference in the fungal glycosylation structures compared to mammalian glycoproteins and, therefore, lower binding efficiency of the lectins. On the other hand, the binding of glycoproteins to lectins is increasing with multivalency. The highly-branched oligomannose structures of glycoproteins in *T. atroviridae* can thus strongly bind to ConA making the elution difficult.

Specific glycoprotein enrichment in human serum samples

As human serum showed a rather more complex elution profile from the lectin beads with a higher number of enriched proteins, further studies were limited to this more comprehensive sample compared to the mycelia sample. Specific glycoprotein enrichment was investigated by an individual incubation of the sample with each individual lectin bead. Against expectations, MCGE electropherograms of all three enrichment experiments showed a very similar peak pattern (**Fig. 5a**). Comparable results were gained in corresponding SDS-PAGE analysis (**ESM Fig. S6a**).

Due to this outcome the interactions between the beads and the complex biological sample were studied in more detail. Therefore, untreated Dynabeads were incubated with human serum according to the established protocol. Resulting SDS-PAGE analysis of the specific elution fraction showed unspecific interactions of the beads with the sample (**ESM Fig. S6b**). Omitting covalent linking of sample constituents to the beads via free binding sites, ethanolamine was used to block all tosyl groups. By this only the surface of the bead can interact with the sample without formation of any covalent links. Again a similar band pattern was observed (**ESM Fig. S6b**). In order to reduce the surface access and thus the interaction of the sample with the bead surface, proteins of different sizes (insulin and cytochrome C) were covalently linked to the beads. However, incubation with human serum resulted again in unspecific enrichment (**Fig. 5b**, **ESM Fig. 6b**). The electropherograms of the specific elution fractions showed a similar peak pattern to the enrichment with ethanolamine blocked beads (**Fig. 5b**). Furthermore, the pattern was comparable to the glycoprotein enrichment using a mixture of lectin beads, however with lower intensities. Only a signal migrating at 33 s (89 kDa) was not detected.

Electrophoretic analyses suggested the enrichment of similar analytes because of comparable band pattern when using dedicated lectin beads and others (insulin, cytochrome C, ethanolamine). However, MS identification of SDS-PAGE separated proteins gave only few identifications for glycoproteins from the gel bands. Mostly carrier proteins, like HSA and immunoglobulins, were identified. This unspecific binding can be expected as mentioned before. On the contrary, for proteins separated by SDS-PAGE after lectin enrichment, mostly

glycoproteins were identified with good significance in more than one analysis (triplicate and quadruplicate analyses of the full method, from glycoprotein enrichment to mass spectrometric identification).

It was concluded that the surface of the beads, which are made from polystyrene with a modified polyurethane layer, most likely has high potential to interact with the complex sample, i.e. the very abundant and sticky proteins. Yet it can be considered that the unspecific interactions of proteins with the beads might be reduced due to shielding effects of the bound lectins of the surface, but it was still recognizable (HSA binding). As a consequence, this information necessarily has to be taken into account for any data interpretation in regard to specificity investigations.

Conclusions

The present study shows the successful combination of lectin affinity enrichment with MCGE for a sensitive and rapid glycoprotein analysis. Results were in very good agreement with corresponding SDS-PAGE findings, but exhibiting as major advantages lower sample amounts and taking lesser time. Additional MS identification proved the selectivity of the method even for complex biological samples. However, the existence of unspecific interactions between analytes (in very complex matrices as serum) and the magnetic beads (particular the evaluated Dynabeads) were observed, which makes certain analyses challenging. A more detailed study by e.g. 2-D gel electrophoresis or HPLC-ESI-MS will be necessary for in-depth studies of lectin selectivities and specificities with magnetic beads. The combination of the bead-based lectin affinity enrichment with MCGE proved to be a strategy opening up new ways in 2-D gel electrophoresis or particular quality control in biotechnology.

Acknowledgements

Financial support was provided in part by an Agilent Technologies - University Relations Program grant. The authors wish to thank Elisabeth Eitenberger (Institute of Chemical Technologies and Analytics, TU Wien, Vienna, Austria) for SEM analyses, Sabine Gruber and Susanne Zeilinger-Migsich (Institute of Microbiology, University of Innsbruck, Innsbruck, Austria) for providing *T. atroviride* mycelia samples, and Sophie Fröhlich (Boehringer Ingelheim, Vienna, Austria) for donating blood samples.

Conflict of Interest

The authors have declared no conflict of interest and Christian Wenz, Susanne Glück, Andreas Rüfer and Martin Kratzmeier are full-time employees of Agilent Technologies.

References

1. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta. 1999;1473:4-8.
2. Varki A. Biological roles of oligosaccharides: all of the theories are correct. Glycobiology. 1993;3:97-130.
3. Dan X, Liu W, Ng TB. Development and applications of lectins as biological tools in biomedical research. Med Res Rev. 2016;36:221-47.
4. Syed P, Gidwani K, Kekki H, Leivo J, Pettersson K, Lamminmaki U. Role of lectin microarrays in cancer diagnosis. Proteomics. 2016;16:1257-65.
5. Mechref Y, Madera M, Novotny MV. Glycoprotein enrichment through lectin affinity techniques. Methods Mol Biol. 2008;424:373-96.
6. Fanayan S, Hincapie M, Hancock WS. Using lectins to harvest the plasma/serum glycoproteome. Electrophoresis. 2012;33:1746-54.

7. Mandal DK, Bhattacharyya L, Koenig SH, Brown RD, 3rd, Oscarson S, Brewer CF. Studies of the binding specificity of concanavalin A. Nature of the extended binding site for asparagine-linked carbohydrates. Biochemistry. 1994;33:1157-62.
8. Schwefel D, Maierhofer C, Beck JG, Seeberger S, Diederichs K, Moller HM et al. Structural basis of multivalent binding to wheat germ agglutinin. J Am Chem Soc. 2010;132:8704-19.
9. Shahidi-Noghabi S, Van Damme EJ, Smagghe G. Carbohydrate-binding activity of the type-2 ribosome-inactivating protein SNA-I from elderberry (*Sambucus nigra*) is a determining factor for its insecticidal activity. Phytochemistry. 2008;69:2972-8.
10. Loo D, Jones A, Hill MM. Lectin magnetic bead array for biomarker discovery. J Proteome Res. 2010;9:5496-500.
11. Yang Z, Hancock WS. Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. J Chromatogr A. 2004;1053:79-88.
12. Bousse L, Mouradian S, Minalla A, Yee H, Williams K, Dubrow R. Protein sizing on a microchip. Anal Chem. 2001;73:1207-12.
13. Kuschel M, Neumann T, Barthmaier P, Kratzmeier M. Use of lab-on-a-chip technology for protein sizing and quantitation. J Biomol Tech. 2002;13:172-8.
14. Wenz C, Marchetti-Deschmann M, Herwig E, Schrottner E, Allmaier G, Trojer L et al. A fluorescent derivatization method of proteins for the detection of low-level impurities by microchip capillary gel electrophoresis. Electrophoresis. 2010;31:611-7.
15. Engel N, Weiss VU, Wenz C, Rufer A, Kratzmeier M, Glück S et al. Challenges of glycoprotein analysis by microchip capillary gel electrophoresis. Electrophoresis. 2015;36:1754-8.
16. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem. 1996;68:850-8.
17. Blum H, Beier H, Gross HJ. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis. 1987;8:93-9.
18. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999;20:3551-67.
19. Mandal DK, Kishore N, Brewer CF. Thermodynamics of lectin-carbohydrate interactions. Titration microcalorimetry measurements of the binding of N-linked carbohydrates and ovalbumin to concanavalin A. Biochemistry. 1994;33:1149-56.
20. Giménez E, Benavente F, Barbosa J, Sanz-Nebot V. Towards a reliable molecular mass determination of intact glycoproteins by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom. 2007;21:2555-63.

21. Sturiale L, Barone R, Palmigiano A, Ndosimao CN, Briones P, Adamowicz M et al. Multiplexed glycoproteomic analysis of glycosylation disorders by sequential yolk immunoglobulins immunoseparation and MALDI-TOF MS. Proteomics. 2008;8:3822-32.
22. Nakano M, Kakehi K, Tsai MH, Lee YC. Detailed structural features of glycan chains derived from alpha1-acid glycoproteins of several different animals: the presence of hypersialylated, O-acetylated sialic acids but not disialyl residues. Glycobiology. 2004;14:431-41.
23. Madera M, Mechref Y, Klouckova I, Novotny MV. Semiautomated high-sensitivity profiling of human blood serum glycoproteins through lectin preconcentration and multidimensional chromatography/tandem mass spectrometry. J Proteome Res. 2006;5:2348-63.
24. Conesa A, Punt PJ, van Luijk N, van den Hondel CAMJJ. The secretion pathway in filamentous fungi: A biotechnological view. Fungal Genet Biol. 2001;33:155-71.
25. Gamauf C, Marchetti M, Kallio J, Puranen T, Vehmaanpera J, Allmaier G et al. Characterization of the bga1-encoded glycoside hydrolase family 35 beta-galactosidase of *Hypocrea jecorina* with galacto-beta-D-galactanase activity. FEBS J. 2007;274:1691-700.

Figure Legends

Fig. 1. Comparison of ConA coupling reactions to tosylactivated Dynabeads and MagSi beads. (**a**) SEM analyses (FEI Quanta 200 with accelerating voltages of 10 and 20 kV, respectively) of both type of beads. (**b**) SDS-PAGE analyses of the supernatant (**S**) and the unspecific elution fraction (**E**) after incubation of the beads with different amounts of Con A (1x: 5.33 µM, 2x: 10.66 µM, 10x: 53.3 µM). (**c**) Amounts of ConA (in pmol and %) in the supernatant and unspecific elution fraction as determined with MCGE (P230 assay).

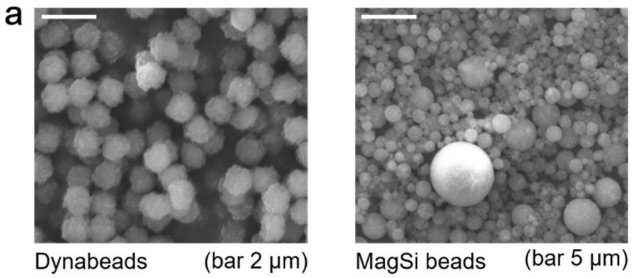
Fig. 2 (**a**) SDS-PAGE and (**b, c**) MCGE (HSP-250 assay) analyses of a Con A enrichment using a mixture of the glycoproteins Tf, A1AT, and AGP with (blue line) and without (black line) β-Gal as negative control. The applied mixtures were analysed before enrichment (**b**) and compared to the specific elution **E** (**c**).

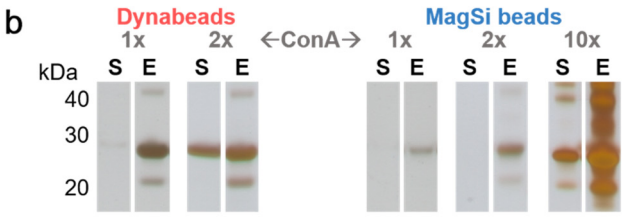
Fig. 3 Glycoprotein enrichment of 2 µL human serum using a mixture of ConA-, SNA-, and WGA-beads (1:1:1; 2 µl each). MCGE analysis (HSP-250 assay) (**a**) before enrichment and (**b**) of the specific elution fraction **E**.

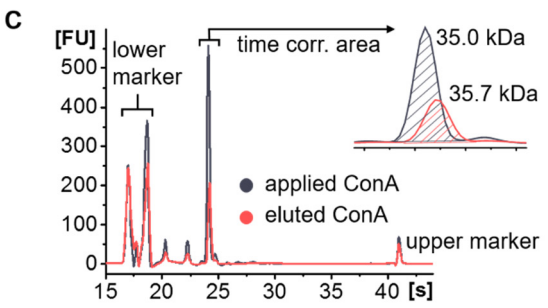
Fig. 4 SDS-PAGE analysis of glycoproteins enriched by a lectin beads mixture (ConA:WGA:SNA 1:1:1, 2 µl each) from human serum (**a**) before and (**b**) after depletion with Pierce Top 12 spin columns, and (**c**) *T. atroviride* using. (**E** – specific elution fraction). Proteins were identified by MALDI-MS after tryptic *in-gel* digestion.

Fig. 5 MCGE analysis (HSP-250 assay) of the specific elution fractions **E** of (**a**) human serum incubated with individual lectin beads in comparison to (**b**) beads coated with ethanolamine, insulin or cytochrome C.

Fig. 1







		applied ConA	eluted ConA
Dynabeads	pmol	53.3	23.9
	%	100	45
MagSi beads	pmol	53.3	2.4
	%	100	5

Fig. 2

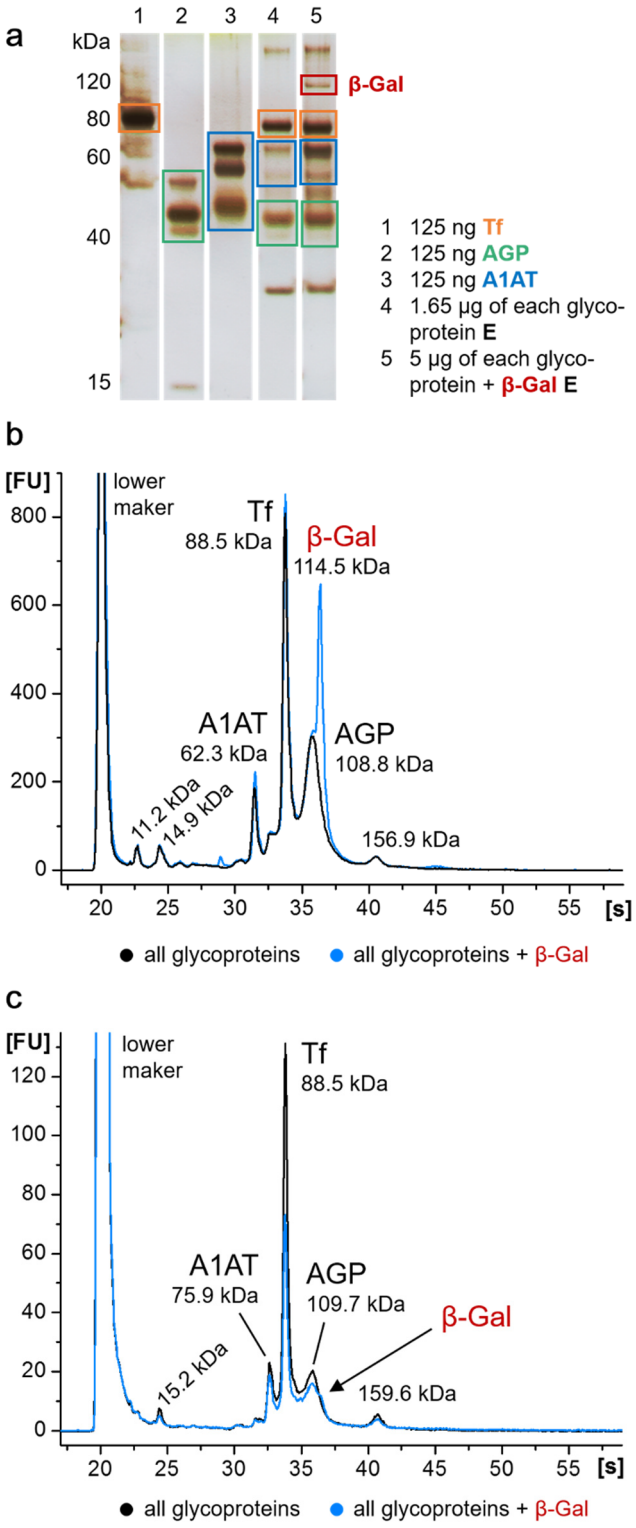

a
1 2 3 4 5
kDa
120
80
60
40
15
β-Gal
1 125 ng Tf
2 125 ng AGP
3 125 ng A1AT
4 1.65 µg of each glyco-protein E
5 5 µg of each glyco-protein + β-Gal E
b
[FU]
800
600
400
200
0
lower maker
Tf
88.5 kDa
β-Gal
114.5 kDa
A1AT
62.3 kDa
AGP
108.8 kDa
11.2 kDa
14.9 kDa
156.9 kDa
20 25 30 35 40 45 50 55 [s]
all glycoproteins
all glycoproteins + β-Gal
c
[FU]
120
100
80
60
40
20
0
lower maker
Tf
88.5 kDa
A1AT
75.9 kDa
AGP
109.7 kDa
β-Gal
15.2 kDa
159.6 kDa
20 25 30 35 40 45 50 55 [s]
all glycoproteins
all glycoproteins + β-Gal

Fig. 3

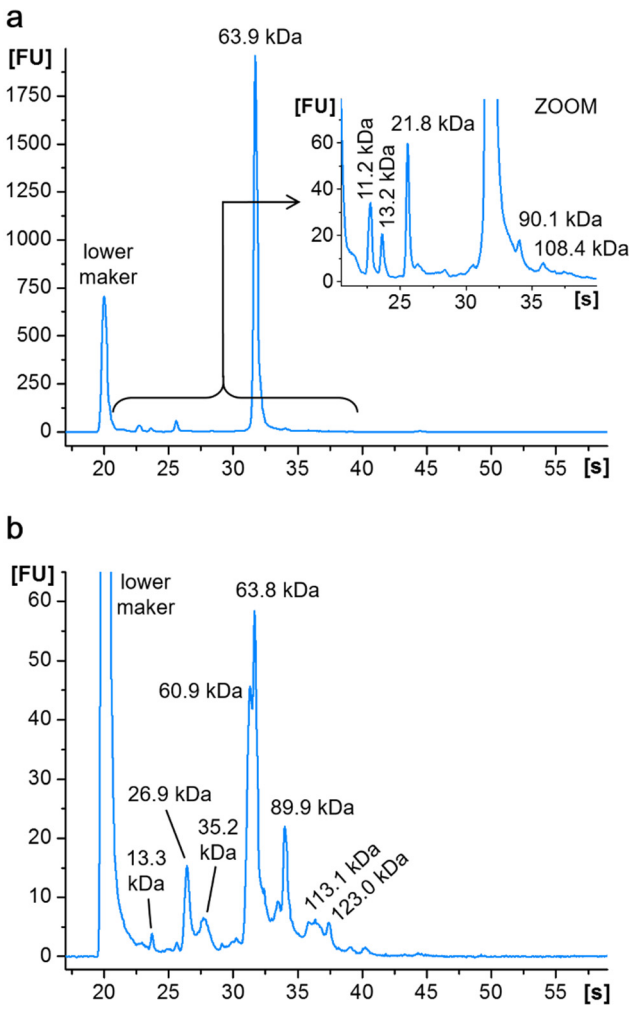
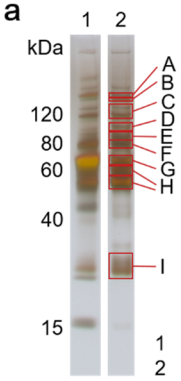

a
[FU]
1750
1500
1250
1000
750
500
250
0
63.9 kDa
lower
maker
[FU]
60
40
20
0
11.2 kDa
13.2 kDa
21.8 kDa
ZOOM
90.1 kDa
108.4 kDa
25
30
35
[s]
20
25
30
35
40
45
50
55
[s]
b
[FU]
60
50
40
30
20
10
0
lower
maker
63.8 kDa
60.9 kDa
26.9 kDa
35.2
kDa
13.3
kDa
89.9 kDa
113.1 kDa
123.0 kDa
20
25
30
35
40
45
50
55
[s]

Fig. 4

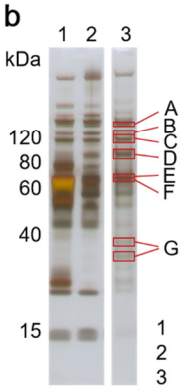
a



	protein	glycosylation sites
A	alpha-2-macroglobulin	8 N-linked
B	complement factor H	6-8 N-linked
C	inter-alpha-trypsin inhibitor heavy chain H4	4 N-linked
	complement C3	3 N-linked
D	plasminogen	1 N-linked, 2 O-linked
E	complement C1r subcomponent	4 N-linked
F	Ig mu chain	5 N-linked
G	serum albumin	-
H	Ig gamma	5 N-linked
I	Ig kappa chain	-

1 1 µl 1:100 serum
2 serum **E**

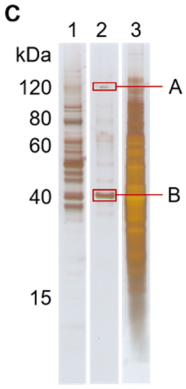
b



	protein	glycosylation sites
A	complement Factor H	6-8 N-linked
B	ceruloplasmin	6 N-linked
C	inter-alpha-trypsin inhibitor heavy chain H4	4 N-linked
	complement C3	3 N-linked
D	plasma protease C1 inhibitor	6-7 N-linked, 7 O-linked
	plasminogen	1 N-linked, 2 O-linked
E	hemopexin	5 N-linked, 2 O-linked
F	complement C3	3 N-linked
G	lectin related	-

1 1 µl 1:100 serum
2 0.3 µl depleted serum
3 depleted serum **E**

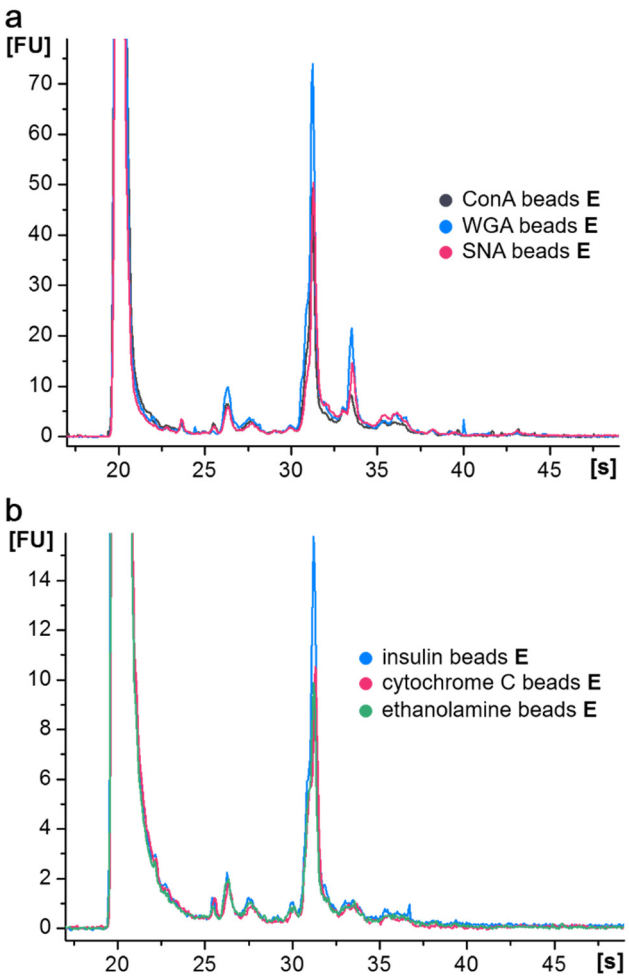
c



1 700 ng pure sample
2 specific elution **E**
3 unspecific elution (SDS)

	protein
A	β-galactosidase
B	glyceraldehyde-3-phosphate dehydrogenase

Fig. 5


a
[FU]
70
60
50
40
30
20
10
0
20
25
30
35
40
45
[s]
ConA beads E
WGA beads E
SNA beads E
b
[FU]
14
12
10
8
6
4
2
0
20
25
30
35
40
45
[s]
insulin beads E
cytochrome C beads E
ethanolamine beads E

Electronic Supplementary Material

Microchip capillary gel electrophoresis combined with lectin affinity enrichment employing magnetic beads for glycoprotein analysis

Nicole Y. Engel[1], Victor U. Weiss[1], Christian Wenz[2], Susanne Glück[2], Andreas Rüfer[2], Martin Kratzmeier[2], Martina Marchetti-Deschmann[1], Günter Allmaier[1]

[1]Institute of Chemical Technologies and Analytics, TU Wien (Vienna University of Technology), Vienna, Austria

[2]Agilent Technologies, Waldbronn, Germany

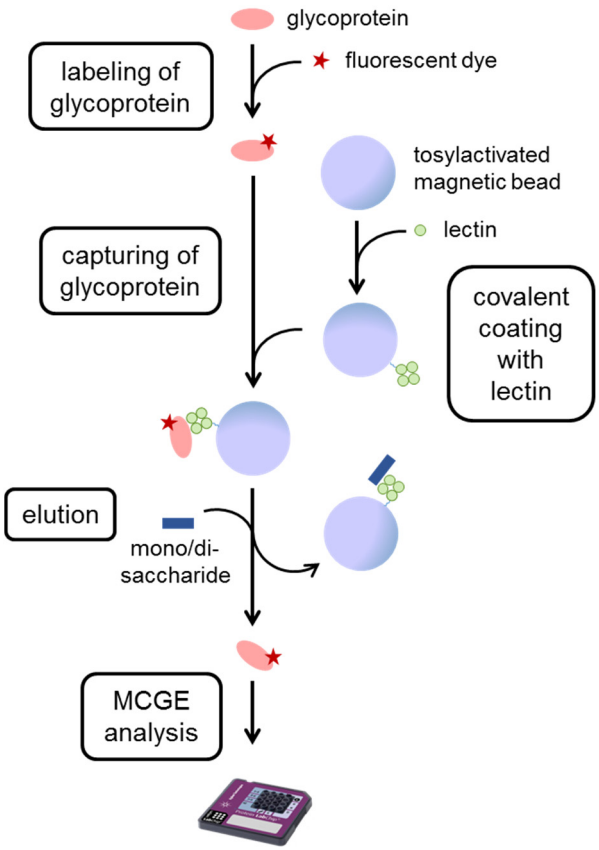
Address reprint requests to:

Prof. Günter Allmaier, Institute of Chemical Technologies and Analytics, TU Wien (Vienna University of Technology), Getreidemarkt 9/164-IAC, A-1060 Vienna, Austria

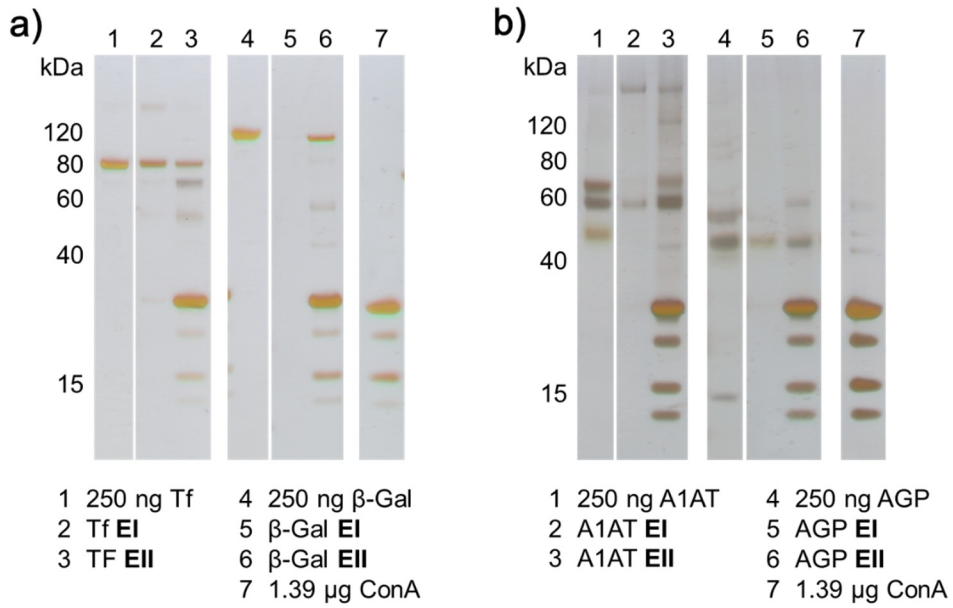
E-mail: guenter.allmaier@tuwien.ac.at

Phone: +43 - 1 - 58801 - 15160

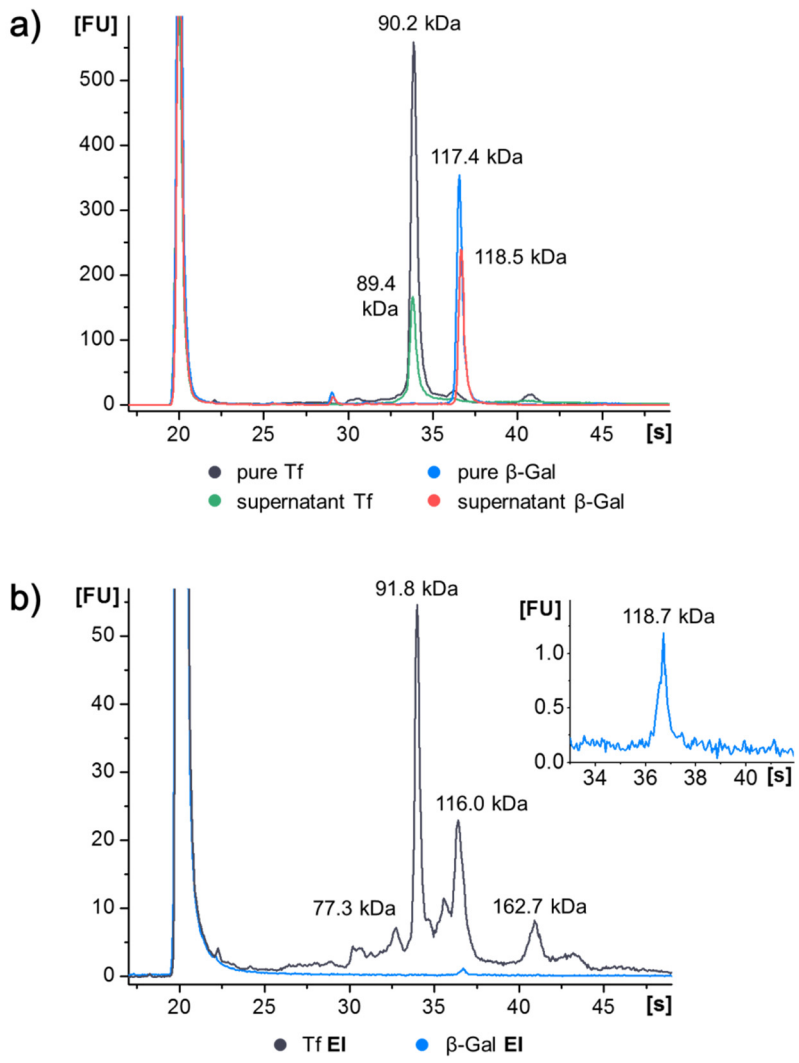
Fax: +43 - 1 - 58801 - 15199



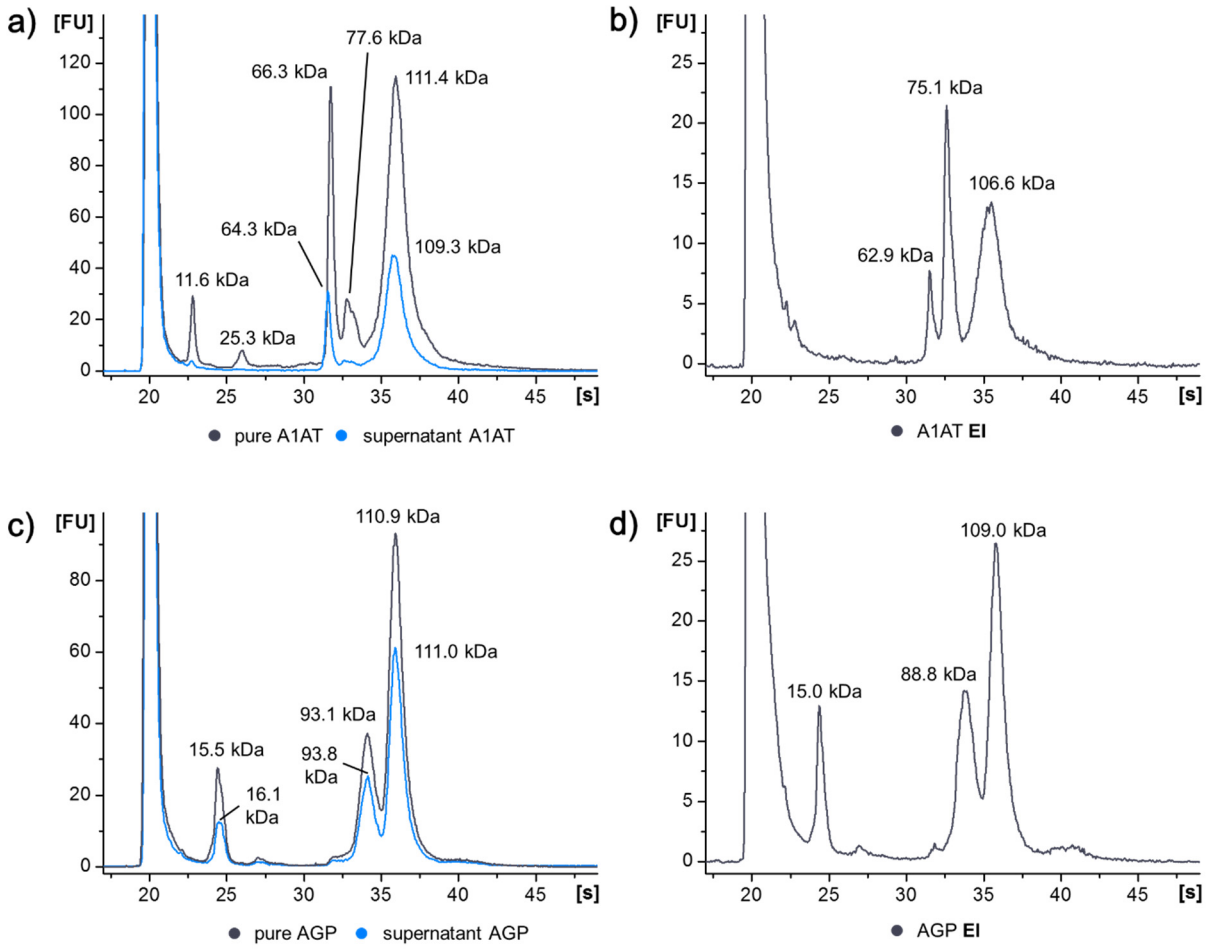
ESM Fig. 1 Workflow of the affinity enrichment of glycoproteins. Lectins were covalently coupled to magnetic beads and incubated with fluorescently labeled analytes. The captured glycoproteins were specifically eluted by the respective complementary mono- and disaccharides and subsequently analysed with MCGE or SDS-PAGE.



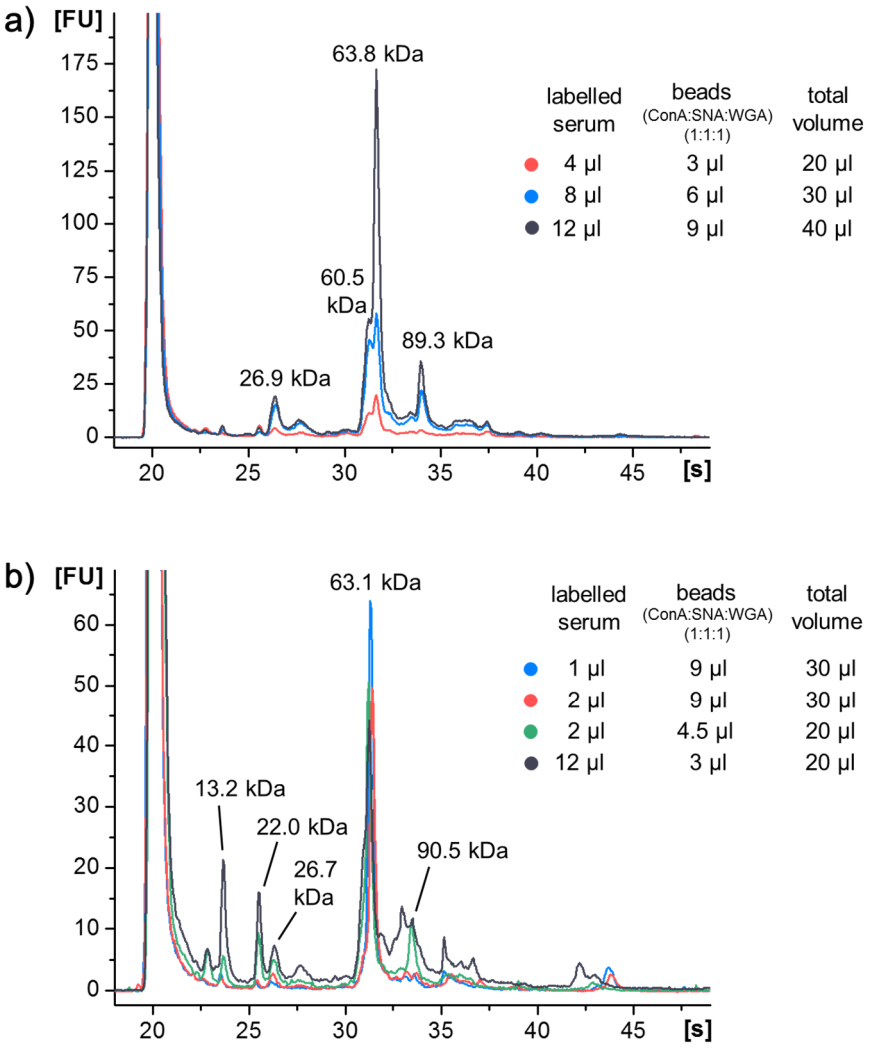
ESM Fig. S2. SDS-PAGE analysis of a ConA enrichment of the glycoproteins (**a**) Tf (lane 1-3), (**b**) A1AT (lane 1-3), and (**b**) AGP (lane 4-6), as well as of (**a**) the non-glycosylated β-Gal (lane 4-6) using ConA-beads. The samples were analysed before (250 ng protein each) and after enrichment (specific elution **EI**, unspecific elution **EII**).



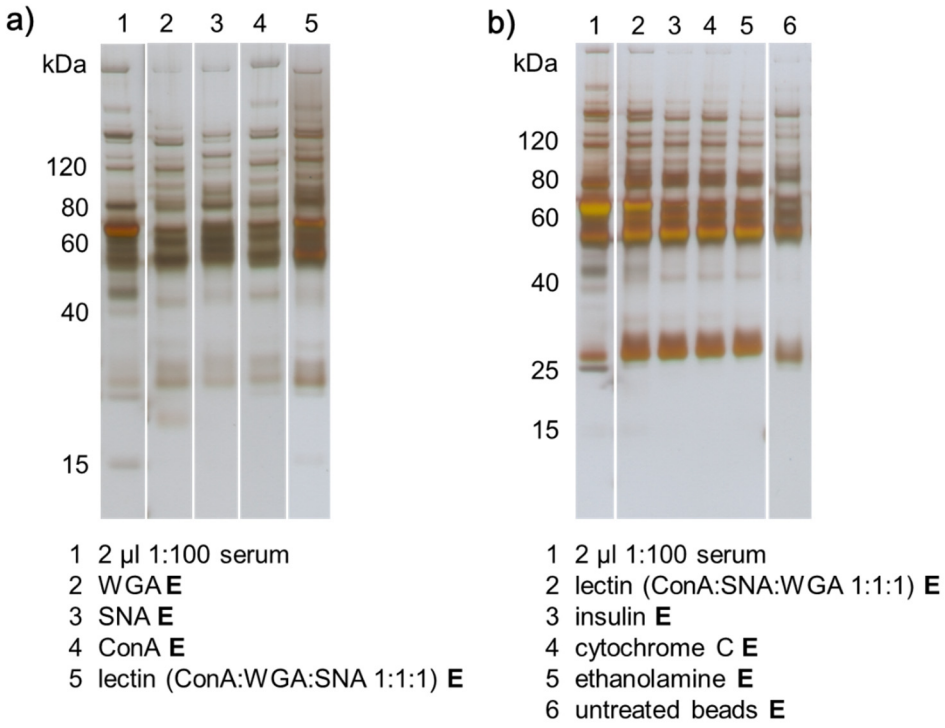
ESM Fig. S3. MCGE (HSP-250 assay) analyses of the ConA enrichments of the glycoprotein Tf and the non-glycosylated β-Gal. The samples were analysed before (**a**: initially applied protein) and after enrichment (**a**: supernatant, **b**: specific elution fraction **EI**).



ESM Fig. S4. MCGE (HSP-250) analyses of ConA enrichments of the glycoproteins (**a**/**b**) A1AT and (**c**/**d**) AGP. The samples were analysed before (**a/c**: initially applied protein) and after enrichment (**a/c**: supernatant, **b/d**: specific elution fraction).



ESM Fig. S5. Enrichment of human serum with a mixture of lectin beads with (**a**) varying serum/bead concentrations or (**b**) serum/bead ratios. Specific elution fractions were analysed by MCGE using the HSP-250 assay.



ESM Fig. S6. Human serum was incubated (**a**) with each lectin bead individually and (**b**) with beads coated with several analytes of different sizes and properties. The specific elution fractions **E** of each enrichment were analysed by SDS-PAGE.

Journal of The American Society for Mass Spectrometry (**2017**), *28,* 77-86

nES GEMMA analysis of lectins and their interactions with glycoproteins – separation, detection, and sampling of non-covalent biospecific complexes

Nicole Y. Engel, Victor U. Weiss, Martina Marchetti-Deschmann, Günter Allmaier

Institute of Chemical Technologies and Analytics, TU Wien, Vienna, Austria

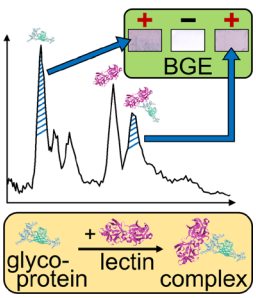
J. Am. Soc. Mass Spectrom. (2017) 28:77–86
DOI: 10.1007/s13361-016-1483-0

FOCUS: 31st ASILOMAR CONFERENCE,
NATIVE MS-BASED STRUCTURAL BIOLOGY: RESEARCH ARTICLE

nES GEMMA Analysis of Lectins and Their Interactions with Glycoproteins – Separation, Detection, and Sampling of Noncovalent Biospecific Complexes

Nicole Y. Engel, Victor U. Weiss, Martina Marchetti-Deschmann, Günter Allmaier

Institute of Chemical Technologies and Analytics, TU Wien (Vienna University of Technology), Getreidemarkt 9/164-IAC, A-1060, Vienna, Austria



Abstract. In order to better understand biological events, lectin–glycoprotein interactions are of interest. The possibility to gather more information than the mere positive or negative response for interactions brought mass spectrometry into the center of many research fields. The presented work shows the potential of a nano-electrospray gas-phase electrophoretic mobility molecular analyzer (nES GEMMA) to detect weak, noncovalent, biospecific interactions besides still unbound glycoproteins and unreacted lectins without prior liquid phase separation. First results for *Sambucus nigra* agglutinin, concanavalin A, and wheat germ agglutinin and their retained noncovalent interactions with glycoproteins in the gas phase are presented. Electrophoretic mobility diameters (EMDs) were obtained by nES GEMMA for all interaction partners correlating very well with molecular masses determined by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) of the individual molecules. Moreover, EMDs measured for the lectin–glycoprotein complexes were in good accordance with theoretically calculated mass values. Special focus was laid on complex formation for different lectin concentrations and binding specificities to evaluate the method with respect to results obtained in the liquid phase. The latter was addressed by capillary electrophoresis on-a-chip (CE-on-a-chip). Of exceptional interest was the fact that the formed complexes could be sampled according to their size onto nitrocellulose membranes after gas-phase separation. Subsequent immunological investigation further proved that the collected complex actually retained its native structure throughout nES GEMMA analysis and sampling.

Keywords: Lectin, Glycoprotein, nES GEMMA, CE-on-a-chip, Electrophoresis

Received: 29 April 2016/Revised: 24 July 2016/Accepted: 13 August 2016/Published Online: 19 September 2016

Introduction

In recent years, the analyses of a variety of macromolecules (e.g. DNA [1], proteins [2–5], polymers [6–8], viruses and virus-like-particles [5, 9, 10], gold nanoparticles [11–13]) have shown the broad applicability of nano-electrospray gas-phase electrophoretic mobility molecular analyzer (nES GEMMA). Thus, this method is used with increasing interest for size-determination of particles ranging from small analytes of only a few nm in size up to particles of several hundred nm. Characterizations with nES GEMMA are generally independent of the analyzed particle type and chemical composition, which makes the method very versatile.

As previously described by Kaufman in detail [3], nES GEMMA separates analytes according to their electrophoretic mobility diameter (EMD) in the gas phase, which can directly be correlated to the dry particle diameters in the nm range. Consequently, the molecular weights can be calculated by application of a correlation derived from respective standard compounds [3, 4]. In brief, multiply charged droplets produced in cone jet mode in the nES unit are dried and simultaneously charge-reduced in a bipolar atmosphere (induced by a ^{210}Po source) and subsequently introduced into the nano differential mobility analyzer (nDMA). Dominantly singly charged analytes with a certain EMD can exit the nDMA at a particular applied voltage. For detection, the so obtained monodisperse

Electronic supplementary material The online version of this article (doi:10.1007/s13361-016-1483-0) contains supplementary material, which is available to authorized users.

Correspondence to: Günter Allmaier; *e-mail:* guenter.allmaier@tuwien.ac.at

aerosol is directed into a condensation particle counter (CPC), in which supersaturated n-butanol vapor condenses onto the particles. Following nucleation, single particles can be counted by laser light scattering yielding e.g., a number-based particle concentration.

nES GEMMA also allows a size-selective collection of analytes after gas-phase separation for consecutive investigations like microscopic measurements (transmission electron microscopy, TEM; atomic-force microscopy, AFM) or a biological test as an immunologic assay [14–16]. For this purpose, the CPC is replaced by an electrostatic nanoparticle sampler (ENAS). It consists of an electrically grounded sampling chamber that features an electrode in its bottom center. By application of a negative voltage to this electrode, positively charged particles coming from the nDMA are attracted. Consequently, they are sampled onto a substrate (e.g., TEM grid, freshly cleaved mica plate or nitrocellulose (NC) membrane) mounted on top of the electrode. The deposition rate is affected by the flow rate, with which the analytes enter the sampling chamber, by the applied voltage as well as by the particle concentration and charge.

Operating at ambient pressure and with nondenaturing electrolyte solutions, nES GEMMA has proven its strength to preserve noncovalent interactions [5, 17–21]. Therefore, nES GEMMA can be considered an effective technique to study even very fragile biocomplexes like lectin–glycoprotein. Lectins have become a major tool in the fields of glycomics and are applied in many methods for a specific glycoprotein enrichment, glycan characterization or targeted glycoprotein detection. Some of the most commonly used lectins are *Sambucus nigra* agglutinin (SNA), wheat germ agglutinin (WGA), and concanavalin A (ConA), with varying specificities towards different oligosaccharide structures. SNA, a lectin isolated from elder, consists of two subunits, A and B, linked by disulfide bridges: the A subunit compromises a N-glycosidase activity, whereas the B subunit is responsible for sugar recognition and binding. The lectin specifically recognizes Neu5Acα(2,6)Gal/GalNAc, sialic acids (N-acetylneuraminic acid Neu5Ac) α-glycosidically linked to galactose (Gal), or N-acetylgalactosamine (GalNAc). It features at least two saccharide-binding sites per B subunit [22]. In comparison, the 36 kDa homodimeric WGA preferably binds to terminal N-acetyl-D-glucosamine (GlcNAc) and its β(1,4)-linked oligomers, as well as to Neu5Ac based on its structural similarity towards GlcNAc. WGA, a plant lectin enriched in the seeds of *Triticum vulgaris*, exhibits four sugar binding sites per monomer [23]. The dimeric form is stabilized by ion pairs, several strong H-bonds, and numerous van der Waals' contacts. The third lectin, ConA, isolated from jack bean (*Canavalia ensiformis*), exists as an oligomer of identical 26 kDa subunits (the exact composition is pH-dependent, see Results and Discussion). It provides one carbohydrate binding site per monomer, which is like the WGA dimer noncovalently linked. ConA specifically binds to mannose (Man) residues as found in the core structure of all N-glycans (Man-α(1,3)[Man-α(1,6)]Man), as well as in high-mannose and hybrid type N-glycans [24, 25].

In the present study, those three lectins were used to analyze their interactions with glycoproteins exhibiting varying glycosylation patterns and degrees for the first time with nES GEMMA. The instrument's advantage of keeping fragile noncovalent biocomplexes intact allowed the separation and detection of the lectin–glycoprotein complexes. It even enabled an investigation of the lectins' binding specificities towards the different applied glycoproteins transferrin (Tf), antitrypsin (A1AT), and acid glycoprotein (AGP), especially in comparison to a nonglycosylated negative control β-galactosidase (β-Gal). The chosen set of glycoproteins differed significantly in size, glycosylation degree, and glycosylation pattern (Table 1): Tf, the biggest of the applied glycoproteins in size, featured the lowest glycosylation content with one O-glycan, two N-glycans, and low degree of sialylation [26]. The smaller A1AT exhibited one additional N-glycosylation site and higher degree of sialylation [28]. AGP was the smallest applied glycoprotein with the highest glycan content (five N-glycans) and the highest number of sialic acid residues attached [30].

It was found that nES GEMMA is a straight-forward method with simplified data interpretation due to charge-reduction to singly charged species compared with ESI mass spectra. Biospecific complexes were detected and, furthermore, sampled onto a NC membrane after gas-phase size-separation in the nDMA for analysis with an immunoassay. The transfer of intact noncovalent complexes to the gas phase was additionally underscored by comparing gained nES GEMMA data with theoretical estimated values based on mass calculations. For several lectins and glycoproteins, molecular masses were measured by matrix-assisted laser desorption/ionization time-of-flight MS (MALDI-TOF-MS) in linear mode. They were in good agreement compared with nES GEMMA-based results demonstrating the applicability of this approach. Owing to the weak interactions, the molecular masses of the biospecific complexes were only determined by nES GEMMA. Lectin–glycoprotein complexes at 10.85 nm diameter (229 kDa) were detected for Tf-SNA and discussed in detail. nES GEMMA-based molecular mass values correlated well with the theoretically calculated masses of the biospecific complexes. Finally, the results of the binding experiments were further confirmed by capillary electrophoresis on a chip (CE-on-a-chip) with laser-induced fluorescence (LIF) detection.

Experimental

Materials

Ammonium acetate (NH_4OAc, ≥99.99%), Tween 20 (bioxtra grade), *N,N*-dimethylformamide, trifluoroacetic acid (TFA, ≥99%), sinapic acid (SA, ≥98%), alkaline phosphatase linked antibody (goat, anti-rabbit immunoglobulin), anti-α_1-antitrypsin antibody (rabbit), and ammonium hydroxide (28.2% ammonia in water) were purchased from Sigma-Aldrich (St. Louis, MO, USA), as were human serum Tf (≥98%), bovine AGP (99%), human A1AT (salt free, lyophilized powder), and β-Gal (lyophilized powder). Lectins SNA,

Table 1. Analysis of Tf [26, 27], A1AT [28, 29], AGP [30], β-Gal [31, 32], and SNA [22, 33] by MALDI-MS and nES GEMMA

Protein	Approx. N-glycosylation (w/w %)[a]	N-glycosylation sites[a]	MALDI-MS MW_{lit} (kDa)[a]	MALDI-MS MW_{exp} (kDa)[b]	nES GEMMA EMD_{exp} (nm)[b]	nES GEMMA MW_{exp} (kDa)[c]	nES GEMMA FWHM (nm)[d]
Tf	6	Asn^{413}, Asn^{611}	80	**79.1** ± 0.1	**7.69** ± 0.04	83.4 ± 1.1	0.31 ± 0.01
A1AT	13	Asn^{46}, Asn^{83}, Asn^{247}		**34.4** ± 0.6	**5.81** ± 0.02	37.7 ± 0.5	0.34 ± 0.01
			51	50.8 ± 0.3	6.58 ± 0.07	53.6 ± 1.6	
AGP	37	Asn^{16}, Asn^{39}, Asn^{76}, Asn^{86}, Asn^{118}	33.8	**31.2** ± 0.5	**5.59** ± 0.05	33.8 ± 0.9	0.34 ± 0.02
			-	45.5 ± 0.3	6.62 ± 0.05	54.5 ± 1.1	
			-	76.0 ± 0.5	7.83 ± 0.04	87.9 ± 1.1	
β-Gal	0	-	116.3	**116.4** ± 0.1	9.35 ± 0.00	147.2 ± 0.0	
			-	Not detectable	**13.35** ± 0.06	429.4 ± 5.7	0.45 ± 0.06
SNA-I[e] $[A\text{-}s\text{-}s\text{-}B]_2$	5	8 putative	A: 33 [f] B: 35[f]	**130.1** ± 0.7	**9.40** ± 0.09	149.6 ± 4.4	0.53 ± 0.10
SNA-I[e] $[A\text{-}s\text{-}s\text{-}B]_4$	10	16 putative	-	Not detectable	11.66 ± 0.12	284.7 ± 8.6	

[a] Values according to references
[b] Dominating (glyco)protein species in bold
[c] Values calculated according to [4]
[d] Calculated after normalization to most abundant peak
[e] A and B represent the subunits of SNA, -s-s- a disulfide bond, and $[\ ]_{2/4}$ a dimeric/tetrameric complex
[f] Determined by SDS-PAGE under reducing conditions

ConA, and WGA were from Vector Laboratories (Burlingame, CA, USA). Sodium chloride (NaCl, ≥99.5%), sodium hydroxide (≥99%), as well as acetonitrile (ACN), hydrochloric acid, magnesium chloride hexahydrate, sodium hydrogen carbonate, tris(hydroxymethyl)aminoethane (Tris), and acetic acid (all analytical grade) were obtained from Merck (Darmstadt, Germany). 5-Bromo-4-chloro-3-indolyl phosphate (BCIP), nitro blue tetrazolium (NBT), and pure nitrocellulose membrane (pore size 0.45 µm) were purchased from Bio-Rad Laboratories (Hercules, CA, USA). Boric acid (pro analysis) and dimethyl sulfoxide (DMSO, pro analysis) were from Fluka (Buchs, Switzerland). Dy-649P1 NHS-ester ($\lambda_{ex/em}$ = 655/676 nm in ethanol according to the manufacturer) for fluorescence (FL) labeling was obtained from Dyomics (Jena, Germany). A 2.5 mM stock solution of the dye in DMSO was prepared for labeling. Further dilutions of the dye were performed applying only DMSO. For all solutions, water of Millipore grade (18.2 MΩcm resistivity at 25 °C) from a Simplicity UV water purification system (Millipore, Molsheim, France) was used throughout the entire investigation. Prior to application, all electrolytes were filtered with 0.2 µm pore size syringe filters (sterile, surfactant-free cellulose acetate membrane; Sartorius, Goettingen, Germany).

Buffers and Sample Preparation

For nES GEMMA analysis, lectins and glycoproteins were dissolved in 20 mM NH_4OAc pH 4.8 or 7.4 adjusted with acetic acid or ammonium hydroxide, respectively. Owing to the requirement of removal of nonvolatile salts (ConA, A1AT, and β-Gal solutions) 10 kDa cutoff spin filters (polyethersulfone (PES) membrane; VWR, Vienna, Austria) were used according to the manufacturer's protocol. All analytes (direct solution or retentate) were then diluted to the required concentration (5–320 µg/mL). They were measured either directly or after 1 h incubation at 24 °C and 650 rpm for interaction experiments.

In the case of CE-on-a-chip experiments, analytes had to be FL labeled prior to electrophoresis. Thus, 150 µg protein (15 µg in the case of β-Gal) in 100 mM sodium borate pH 8.3 were mixed with 5 µM dye and incubated overnight in the dark at room temperature. Nonreacted dye was subsequently removed in the same way as described for the desalting step. Analyte concentrations were adjusted to 50–250 µg/mL with sodium borate prior to analysis. Analytes were either measured directly or after 1 h incubation of lectin and glycoprotein at 24 °C.

nES GEMMA

nES GEMMA experiments were carried out on a system consisting of a model 3480 electrospray aerosol generator including a ^{210}Po source, a model 3080 electrostatic classifier containing a nDMA unit, and a n-butanol driven model 3025A ultrafine CPC from TSI Inc. (Shoreview, MN, USA).

For operation in detection mode, the nDMA sheath flow was set to 15 liters per minute (Lpm; particle separation size range 2.0–64.4 nm EMD), for sampling a flow of 14 Lpm (2.0–67.3 nm EMD) was used. Samples were introduced via a 25 cm long cone-tipped fused silica capillary with an inner and outer diameter of 40 and 150 µm, respectively; 4 psid (pounds per square inch differential, approximately 0.3 bar) of pressure were applied to the sample vial for analyte introduction to the nES capillary in detection mode, whereas 2 psid were used for sampling. Higher pressure during long sampling experiments destabilized the spraying process and was thus avoided. The nES sheath gas (CO_2 and filtered, dried air from a membrane dryer Superplus, Ludvik Industriegeräte, Vienna, Austria) was set to 0.6 Lpm and voltages were adjusted for a stable cone jet

mode (2.0–2.5 kV). A median of 10 scans, 120 s each (100 s scan time, 20 s retrace time), yielded a spectrum (as shown in figures) and was used for data interpretation with the OriginPro software (v 9.1.0, OriginLab, Northampton, MA, USA).

For size-selected particle collections, a 3089 ENAS (TSI Inc.) replaced the CPC. The NC membrane was cut to 15 mm square. It was mounted on top of the center electrode using double-sided adhesive tape (Scotch/3 M, St. Paul, MN, USA), which was removed after sampling. The ENAS was operated at –9.5 kV and a gas flow rate of 1 Lpm. During collections of three times 12 h on three consecutive days about 475 µL of sample volume (20 µg/mL A1AT, a mixture of 10 and 20 µg/mL A1AT and SNA, respectively, or pure 20 mM NH_4OAc, pH 7.4, as blank) were consumed.

Capillary Electrophoresis-on-a-Chip

CE-on-a-chip was carried out on an Agilent 2100 Bioanalyzer platform (Waldbronn, Germany), a chip-based microfluidic system based on LIF detection (red diode laser, $\lambda_{ex/em}$ = 635/685 nm). Owing to software modifications, large particles (e.g., viruses, protein complexes) are separated on chips originally designed for nucleic acid separation according to charge and size as previously described [34]. Briefly, the microfluidic channels were filled with 100 mM sodium borate pH 8.3 as background electrolyte (BGE) by applying pressure for 20 s using the Agilent Chip Priming Station. Twelve µL of BGE each were then applied to waste and buffer wells, and 6 µL of labeled analyte solutions to the sample wells. Prior to sample analysis on each chip, the separation channel was electrophoretically flushed with dye solution, 12.5 nM Dy-649P1 in BGE, followed by setup of the instruments optics and electrophoretic removal of the dye. Data were collected via the red laser of the instrument with the Agilent 2100 Expert software, exported, and plotted using the OriginPro software.

Immunological Assay

After sample collection with nES GEMMA onto 0.45 µm NC membrane, the substrate was removed from the ENAS and tested for the presence of antitrypsin. Simultaneously, 10 and 50 ng of antitrypsin and SNA were directly applied to a control membrane and examined under the same conditions (dot blot assay). Both substrates were washed with TBS-Tween (20 mM Tris pH 8.3, 154 mM NaCl, 0.1% Tween 20) for 30 min and incubated overnight with anti-α_1-antitrypsin antibody (1:9000, v:v, in TBS-Tween). They were washed three times in TBS-Tween for 5 min each, followed by an incubation with the anti-rabbit antibody conjugated to alkaline phosphatase (1:10,000, v:v, in TBS-Tween). The washing steps were repeated and the membranes further washed with TBS without Tween for 5 min. For color visualization, BCIP and NBT were prepared according to the manufacturer's instructions and used for a 15 min incubation step. The reaction was stopped by addition of water.

MALDI-MS

Experiments were performed on the MALDI-TOF-MS AXIMA TOF^2 and, in the case of β-Gal, on the AXIMA-CFR *plus* instrument (Shimadzu Kratos Analytical, Manchester, UK) both equipped with nitrogen laser (λ = 337 nm). Both instruments were operated in linear positive ion mode. Samples were prepared on stainless steel MALDI target plates using the dried-droplet technique. Glycoprotein and β-Gal samples were applied 1:1 (v:v) ratio with 10 mg/mL SA in 0.1% TFA/ACN (1:1, v:v) as MALDI-MS matrix to a final amount of 10–20 and 1.5 pmol, respectively, on target and dried at room temperature.

Results and Discussion

Individual nES GEMMA Analysis of Glycoproteins and Lectins

For determination of the EMD, each analyte was measured individually with nES GEMMA at different concentrations in 20 mM NH_4OAc (pH 7.4). The chosen buffer system should (1) be volatile, (2) resemble physiological conditions for glycoprotein–lectin interactions, and (3) be appropriate for a stable electrospray process. For experiments including lectins, higher NH_4OAc concentrations destabilized the Taylor cone at the nES capillary tip and were consequently avoided.

Figure 1 exemplarily displays the nES GEMMA spectra of the lectin SNA, the glycoprotein AGP, and the nonglycosylated protein β-Gal employed as negative control. For nES GEMMA spectra of the glycoproteins A1AT and Tf, as well as of the lectins WGA and SNA refer to the Supplementary Information. Figure 1a shows a dominating singly charged peak $[2\ M]^+$ representing a dimer of SNA with an EMD of 9.40 ± 0.09 nm, which corresponds to a MW of 149.6 ± 4.4 kDa calculated from an EMD/MW correlation [4]. This value is slightly deviating from the MALDI-MS derived MW of 130.1 ± 0.7 kDa (Table 1). SNA consists of four subunits (two of each identical; 2AB) held together by intramolecular disulfide bridges [35]. Owing to structure flexibilities of this complex in gas phase, the protein might appear bigger in nES GEMMA experiments with a higher MW calculated than measured with MALDI MS. Additionally, the singly charged tetramer $[4\ M]^+$ can be observed, which is especially apparent at higher concentrations. With increasing concentrations more than one analyte can be statistically present in a sprayed droplet, which leads to the formation of nonspecific gas-phase singly charged oligomers formed during the nES process [2]. These artificial oligomers can be distinguished from naturally formed biospecific complexes by a rapid loss of signal intensity or even disappearance with decreasing concentrations. Yet, lectins have a high tendency to aggregate. The fact that the tetramer signal did not completely vanish even at low concentrations points to biologically relevant tetramer formation already in solution.

In contrast, oligomer formation in the case of glycoproteins AGP (Figure 1b) was merely concentration-dependent and,

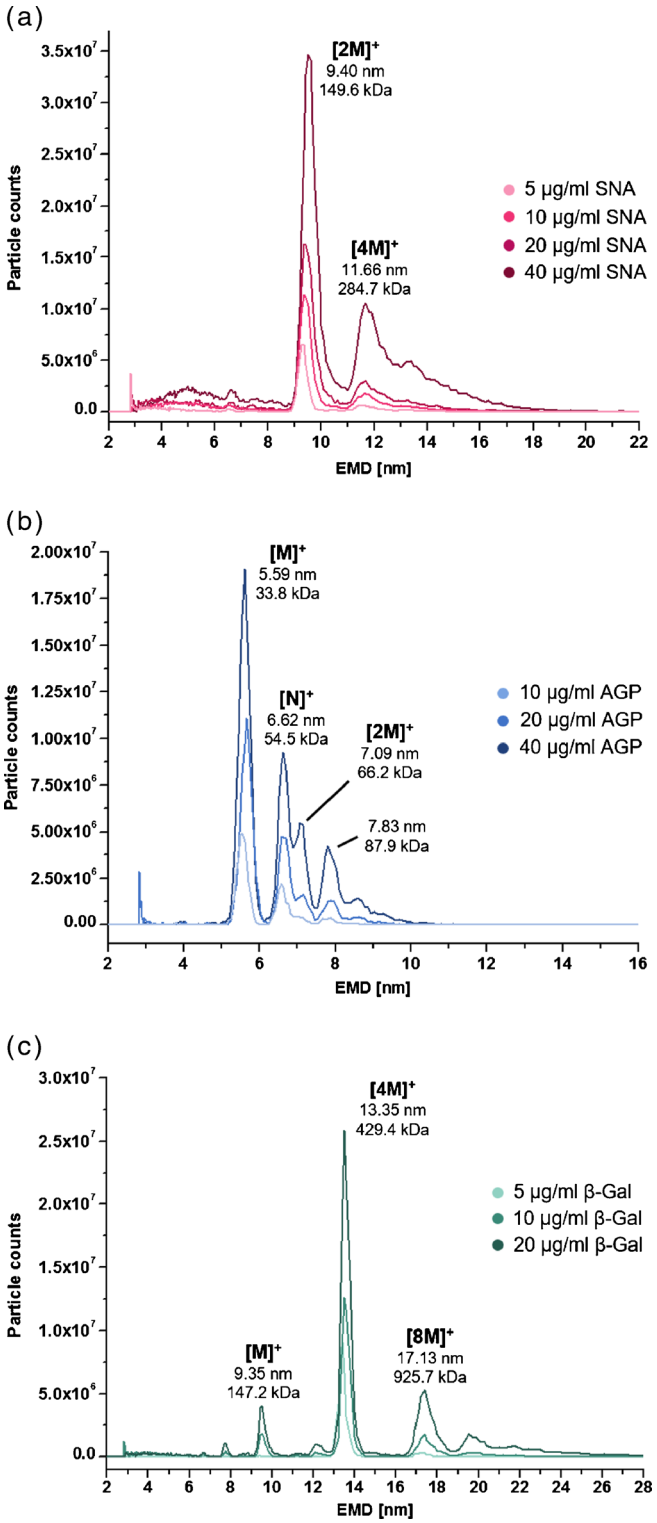


Figure 1. nES GEMMA analysis of different concentrations of the lectin SNA **(a)**, the glycoprotein AGP **(b)**, and the nonglycosylated β-Gal **(c)**. The subunits A and B of SNA are presented as M (M = AB) **(a)**. $[N]^+$ represents a second constituent of AGP **(b)**

hence, nES-induced. Furthermore, the existence of several AGP species with the most abundant one at 5.59 ± 0.05 nm (33.8 ± 0.9 kDa) was confirmed. These results were in good accordance to MALDI-MS data having, however, slightly higher values. Tf showed also gas-phase oligomerization (Supplementary Figure 1a) and A1AT likewise consisted of several species (Supplementary Figure 1b). β-Gal, on the other hand, a tetramer consisting of four identical, noncovalently linked 116 kDa subunits [31], showed only a less intensive peak of the monomer (9.35 ± 0.00 nm, calculated 147.2 ± 0.0 kDa). The detection of a high abundant tetrameric species (13.35 ± 0.06 nm, calculated 429.4 ± 5.7 kDa) demonstrated the instrument's ability to keep noncovalent oligomers intact during analysis. In comparison, MALDI-MS revealed a MW of the monomeric species of 116.4 ± 0.1 kDa with the applied matrix. Table 1 summarizes the data for all investigated (glycol)proteins and lectins.

In accordance to previously published data on glycoprotein analysis by microchip capillary gel electrophoresis (MCGE) and SDS-PAGE [36], increasing glycan content led to signal broadening in MALDI-MS (Supplementary Figure S2). In contrast, the degree of glycosylation did not affect peak width or MW determination for nES GEMMA. This is in favor of analysis, high reproducibility of EMD values with deviations ≤ 1%, and small peak width (FWHM below 0.34 nm for all glycoproteins). However, this fact can also be considered as disadvantageous in regard to a loss of information about the glycosylation degree itself. Instead, FWHM values of peaks from nES GEMMA spectra were rather influenced by increasing EMDs (Table 1). In summary, gas-phase electrophoresis offers to be a reliable (±1%–5% mass accuracy for 8 kDa–1MDa proteins and protein complexes; reproducibility mostly better than ±0.1 nm) [37], sensitive (attomole amounts total consumption) [2], and fast (120 s per scan) alternative for glycoprotein analysis.

The nES GEMMA spectra of the other two lectins, WGA und ConA, in contrast, were more complex. The WGA spectrum was composed of several components, and ConA showed the formation of many oligomers (Supplementary Figure S1). In addition, the latter proved itself to be rather challenging during analysis because its high degree of oligomerization contributed to capillary clogging. This oligomerization is known to be pH sensitive: at ≤ pH 5 the lectin forms predominantly dimers and at pH ≥ 7 it primarily exists as tetramer [25]. This could also be shown by nES GEMMA (Figure 2a). At pH 4.8 mostly the dimeric form with only a small amount of tetrameric species could be observed. Those ratios were reverted at physiological pH. Next to ConA only β-Gal was affected by pH, which was unstable and not measurable from the acidic electrolyte.

Also known from literature is the fact that ConA requires the presence of divalent cations, e.g., calcium (Ca^{2+}), for correct folding and carbohydrate recognition [25]. However, since high salt concentrations can lead to uncontrolled cluster formation in nES GEMMA [38], different $CaCl_2$ concentrations were tested. No interferences were detected up to 10 µM $CaCl_2$ in NH_4OAc at pH 7.4 (data not shown). The salt addition stabilized the formation of biologically dominant ConA tetramers at low lectin concentrations and was thus considered as

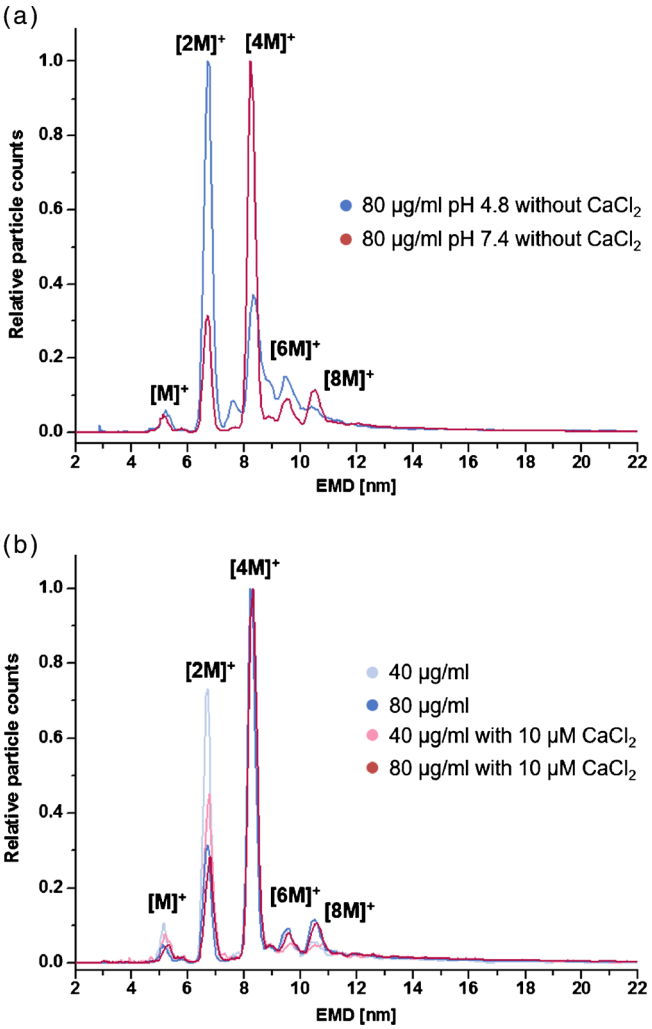


Figure 2. nES GEMMA analysis of the lectin ConA at different pH values **(a)** and at pH 7.4 with addition of 10 µM $CaCl_2$ **(b)** in regard to oligomerization. ConA tetramers [4 M]$^+$ are the biological dominant form

appropriate for ConA interaction studies with glycoproteins (Figure 2b). At higher $CaCl_2$ concentrations, measurements of ConA were not feasible and, therefore, an additional influence of $CaCl_2$ not investigable.

nES GEMMA Interaction Analysis of the Lectins with Glycoproteins

In order to investigate the interaction of SNA, ConA, and WGA with the glycoproteins, the lectins were incubated with each glycoprotein separately at different concentrations and subsequently analyzed with nES GEMMA. Additionally, experiments were carried out with β-Gal as a nonglycosylated negative control.

Owing to the fact that all in this study using glycoproteins showed various degrees of sialylation, a recognition by SNA was expected to be positive in all cases but with different affinities and, i.e., various intensities. Keeping the glycoprotein concentration constant during measurements and increasing only the amount of lectin, a steady decrease of the glycoprotein signal hints the formation of the biospecific complex with SNA. The emerging complex is expected to be detected at the respective EMD ($EMD_{calculated}$), which can be calculated from the sum of the individual MWs and the given EMD / MW correlation [4]. Furthermore, data ($EMD/MW_{experimental}$) can be compared with theoretical values for the $MW_{calculated}$ of the lectin–glycoprotein complex. A close agreement of both values confirms the detection of the non-covalent complex.

Figure 3a presents the incubation of SNA with AGP, which has the highest degree of sialylation. As expected, the intensity of the monomeric AGP signal at 5.55 nm decreased by 75% with increasing SNA concentrations. Moreover, the biospecific complex at 10.06 nm EMD could clearly be detected. In comparison, no according signals were observed for interactions of SNA with the nonglycosylated β-Gal (complex expected at 14.76 nm EMD, Figure 3b). This proved for the first time the capability of nES GEMMA to detect specific lectin-glycoprotein bindings, bindings that are rather weak and, therefore, difficult to analyze (dissociation constants in the mM to high nM range, antibody-epitope bindings are 100- to 1000-fold stronger).

Similar results as with AGP could be gained during the incubations of SNA and A1AT (Supplementary Figure S3a). For A1AT also the SNA concentration was kept constant while steadily increasing the amount of A1AT. Results were the same; the expected signal of the noncovalent complex was observed while the SNA peak decreased (Supplementary Figure S3b). The analysis of the interaction of Tf with the lectin SNA led to comparable findings (Supplementary Figure S3c). However, contrary to AGP and A1AT, the signal for the complex was not as distinct and exhibited lower signal intensities. From this, a lower binding specificity of SNA towards Tf could be concluded, which is in agreement with the comparably lower degree of sialylation. From these findings, we conclude that nES GEMMA can distinguish different lectin binding strengths and specificities towards varying glycoproteins.

The interactions of ConA and WGA with each glycoprotein and β-Gal were additionally investigated to get a more profound understanding of nES GEMMA capabilities (for exemplary results, see Supplementary Figure S4). In the case of ConA, a direct detection of the complex signals was significantly impeded by the lectin's own oligomer peaks, which overlaid the expected glycoprotein–ConA complex. Nevertheless, the decrease of the glycoprotein signals could be observed and used as an indicator for a positive binding: the Tf peak showed the greatest reduction followed by AGP, whereas the A1AT peak diminished only slightly. Also the β-Gal signal decreased slightly, which hinted to minor unspecific interaction between the nonglycosylated protein and ConA.

Investigating glycoprotein interactions with WGA turned out to be rather challenging. Owing to similar MWs of the lectin monomers/oligomers with the glycoproteins, the lectin

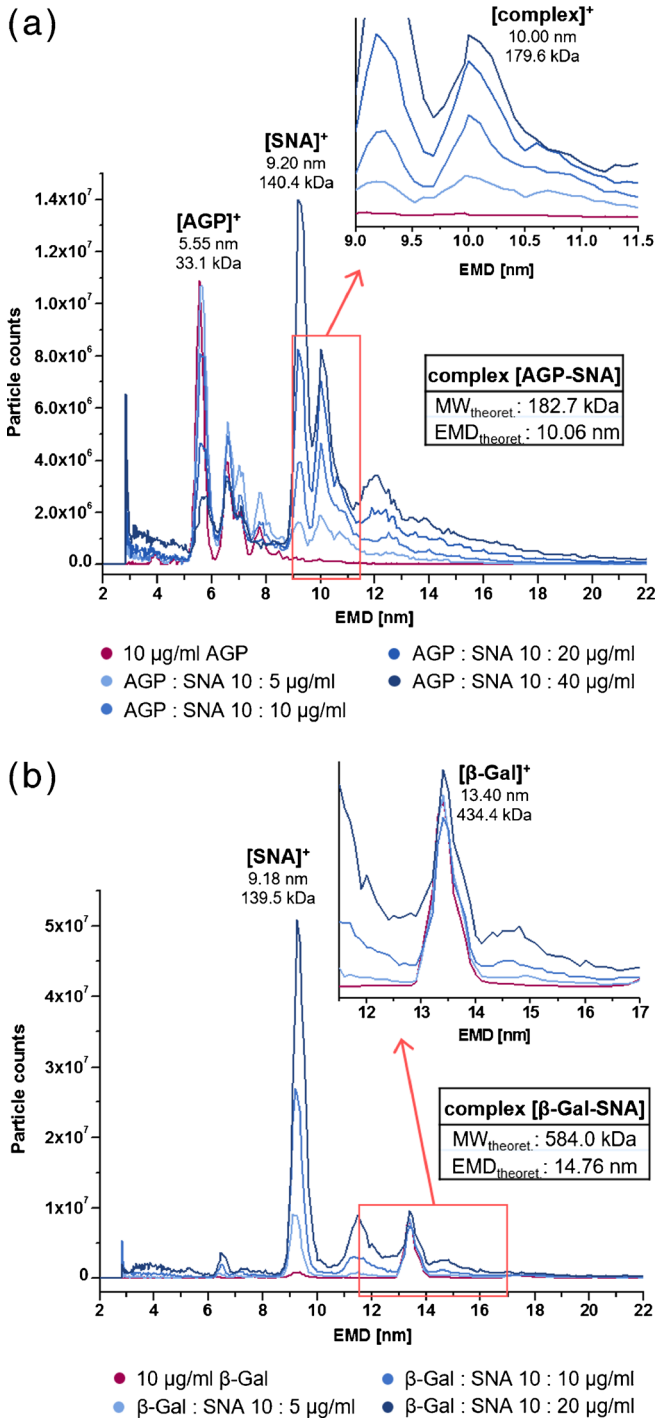


Figure 3. nES GEMMA analysis of AGP **(a)** or β-Gal as negative control **(b)** incubated with different concentrations of SNA

signals did not only overlay the lectin–glycoprotein complex peaks but also those from the glycoproteins. Therefore, neither the decrease in glycoprotein signal nor the newly formed complex signal could be observed. Enhanced resolution is expected for instruments having higher sheath flow rates (e.g., the second generation MacroIMS device from TSI Inc., PDMA [39, 40], or a Vienna type DMA [41]) allowing, then, hopefully for improved signal separation. As a consequence of these findings, additional investigations concentrated on SNA, which showed the most convincing results so far.

Interaction Analysis of SNA by Means of CE-on-a-Chip Experiments

For confirmation of nES GEMMA results, the formation of biospecific lectin–glycoprotein complexes was additionally examined by CE-on-a-chip, a liquid-phase based chip electrophoresis system. Fluorescence labeled glycoproteins and the nonglycosylated β-Gal were incubated with different concentrations of unlabeled SNA. As with nES GEMMA, the formation of a new interaction-relevant signal and the decrease of the glycoprotein peak were expected for rising SNA concentrations. Figure 4a shows the slightly declining signal of AGP with rising SNA content and the clearly emerging

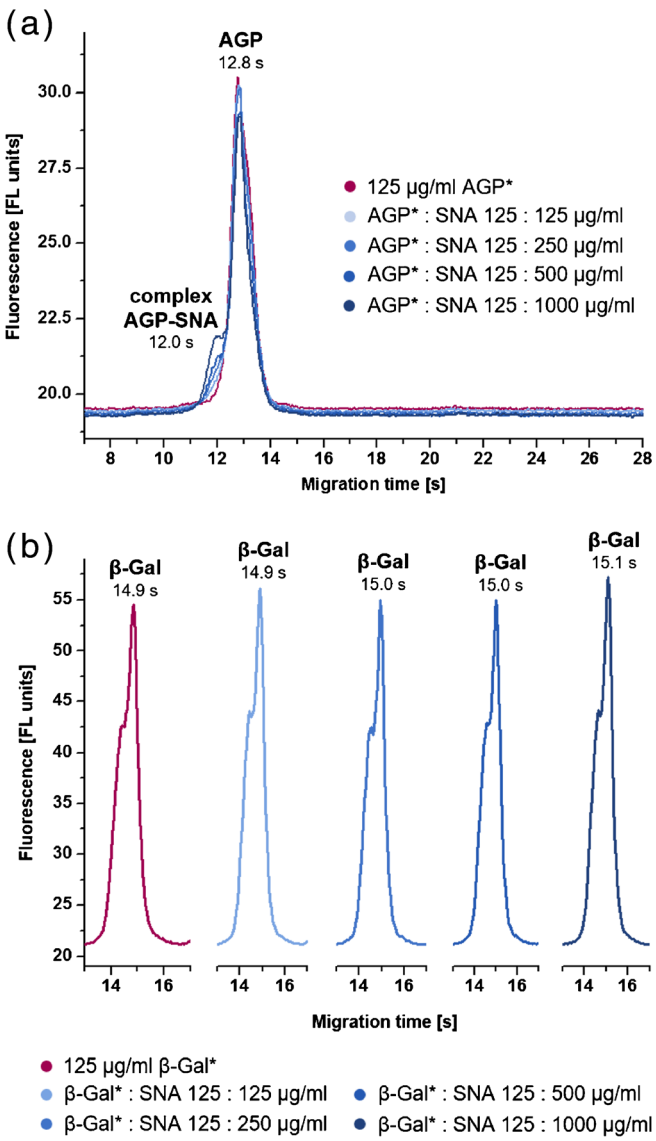


Figure 4. CE-on-a-chip analysis of SNA with AGP and β-Gal: electropherograms of incubations of AGP **(a)** and β-Gal **(b)** with increasing concentrations of unlabeled SNA, respectively. Labeled proteins are marked with an asterisk (*)

glycoprotein-lectin peak at 12.0 s. The negative control β-Gal repeatedly showed no interaction with SNA, maintaining a constant migration pattern despite increasing SNA concentrations (Figure 4b). For A1AT a decrease of signal intensity was observed, whereas the signal for the complex was growing significantly (Supplementary Figure S5a). In addition, it became obvious that the SNA–A1AT complex exhibited the same migration time as a for us today unknown constituent of A1AT (marked with an asterisk in Supplementary Figure 5). The fact that at constant A1AT concentration the signal at 12.6 s showed up to six times increased intensities with rising SNA content allowed for the conclusion that this peak in fact is induced by the glycoprotein–lectin complex. The drastic change in the peak pattern of A1AT hinted a strong interaction with SNA, which was more explicit than with AGP. Tf interacted likewise stronger with SNA than AGP (Supplementary Figure S5b). Thus, all three glycoproteins proved to interact with SNA as already shown with nES GEMMA. Consequently, these experiments corroborated nES GEMMA findings. Reduced or altered binding between AGP and SNA, as detected with CE-on-a-chip, might result from covalently bound FL labels to glycoproteins. They can modify the protein structure and, therefore, influence the binding strength and specificity towards the lectin.

Collection of the Biospecific Lectin–Glycoprotein Complex and Its Immunological Identification

SNA-A1AT complexes were collected after gas-phase size-separation with an ENAS on a NC membrane. After sampling the membrane was removed for subsequent immunologic analysis with colorimetric detection. The color formation on the membrane is based on an epitope recognition of the protein in its native conformation by the antibody. Therefore, it requires the preservation of the collected particles' three-dimensional structure throughout the separation with nES GEMMA and collection process.

By applying A1AT directly on the NC membrane, detection limits for the chosen dot blot assay down to 10 ng glycoprotein were revealed. Based on this, the necessary sampling time of about 36 h was calculated from the applied A1AT-SNA concentrations (10 and 20 ng/µl, respectively, Figure 5a and Supplementary Figure S6) and the injection rates (2 psid of applied pressure). For these 36 h we assumed that (1) less than 5% (usually about 1%) of the overall electrosprayed analytes are reduced to singly charged particles in the neutralizing chamber [42], (2) the sample is a mixture of A1AT, SNA, and A1AT–SNA complex, from which only the latter is of interest for analysis and, therefore, collected onto the NC membrane, (3) that at least 30% to 50% of the present A1AT is forming a complex with SNA, and (4) that no singly charged complex particle is lost during nDMA separation and NC collection. From this we expected about 20 ng glycoprotein–lectin complex to be finally collected on the NC, amounts sufficient for dot blot like analysis.

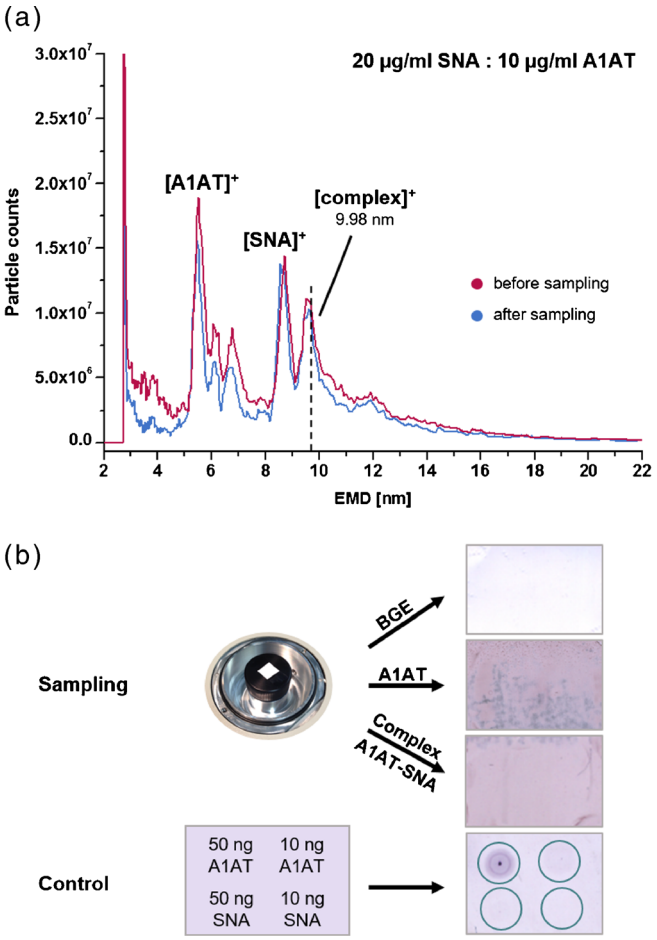


Figure 5. Collection of SNA–A1AT complexes using an ENAS (particle fraction collector). The complex was collected onto NC at 9.96–10.05 nm for 36 h on three consecutive days **(a)** exemplarily showing the sampling of 1 day) followed by immunological identification via color visualization in comparison to a control dot blot experiment **(b)**. For further verification, also pure BGE (9.98 nm) and A1AT (5.60–5.65 nm) were sampled onto NC membrane and immunologically examined **(b)**. The dotted line marks the EMD of sampling of the exemplary day **(a)**

The glycoprotein–lectin complex was sampled at 9.96–10.05 nm EMD, and pure A1AT was collected at 5.60–5.65 nm EMD for immunologic analysis (Figure 5b). Additionally, the BGE was sprayed as a blank for 36 h and sampled at the respective EMDs. In order to verify that the dot blot analysis was specific for A1AT but not SNA or its oligomers, a control was carried out by direct application of SNA and A1AT on NC membranes. Only A1AT showed interaction, proving that any color formation was a direct correlation to A1AT presence. First, the preservation of the native conformation after gas-phase separation of A1AT alone was checked by staining the NC membrane after sampling, which could be observed visually compared with the BGE blank. We found that also the sampling of the SNA–A1AT complex onto the NC membrane showed a noticeable staining comparable to A1AT sample. Interestingly, no distinct spot in the size of the ENAS electrode (9.5 mm diameter) was found, as observed previously after collecting significantly larger particles [16]. In our case,

the applied NC membrane was evenly stained, probably due to the fact that the ENAS voltage was not high enough to deviate the particles from their trajectory imposed by the high nDMA sheath flow and to focus them on a distinct area. An increase of the applied voltage could solve this problem and lead to a shorter sampling time as the analyte concentration would be increased on the NC membrane. However, due to instrument limitations, this approach cannot be realized at the moment.

Conclusions

The nES GEMMA system is a promising platform for the analysis of lectin–glycoprotein interactions as shown in the given study for the first time. Especially, data interpretation is much easier for singly charged particles than ESI spectra of multiply charged noncovalent complexes (data deconvolution can be omitted). Furthermore, sample preparation is only dependent on the formation of a complex in NH_4OAc at the appropriate pH. Today, noncovalent interaction studies are of utmost interest for a better understanding of biological interactions (as for example in molecular machines). nES GEMMA is a valuable tool to study lectin–glycoprotein interactions in regard to interaction specificities and binding strength. We found that the ambient setup of the instrument allowed for the detection of rather weak interactions, which are difficult to maintain in vacuum-based mass analyzers. Working under relatively soft conditions, nES GEMMA even enables sampling of these complexes in their biologically native form. For the first time, bionanoparticles in the rather low nm size range were collected by the ENAS device and analyzed by an immunologic assay. ENAS sampling corroborated correct peak assignment of the noncovalent complex consisting of the lectin SNA and the glycoprotein A1AT in mixed samples. Moreover, it showed the maintenance of interactions formed in liquid phase during gas-phase separation without affecting the native state of the complex. This finally confirms that nES GEMMA allows for separation and detection of biospecific, noncovalent complexes (but relatively weak compared with virus-antibody or virus-like particle-antibody fragment complexes), as well as their successful sampling for further analyses.

Acknowledgments

Open access funding provided by TU (Wien). This project was funded by the Austrian Science Foundation (FWF, Project Number TRP 29-N20).

Open Access

References

1. Mouradian, S., Skogen, J.W., Dorman, F.D., Zarrin, F., Kaufman, S.L., Smith, L.M.: DNA analysis using an electrospray scanning mobility particle sizer. Anal. Chem. **69**, 919–925 (1997)
2. Kaufman, S.L., Skogen, J.W., Dorman, F.D., Zarrin, F., Lewis, K.C.: Macromolecule analysis based on electrophoretic mobility in air: globular proteins. Anal. Chem. **68**, 1895–1904 (1996)
3. Kaufman, S.L.: Analysis of biomolecules using electrospray and nanoparticle methods: the gas-phase electrophoretic mobility molecular analyzer (GEMMA). J. Aerosol Sci. **29**, 537–552 (1998)
4. Bacher, G., Szymanski, W.W., Kaufman, S.L., Zollner, P., Blaas, D., Allmaier, G.: Charge-reduced nano electrospray ionization combined with differential mobility analysis of peptides, proteins, glycoproteins, noncovalent protein complexes and viruses. J. Mass Spectrom. **36**, 1038–1052 (2001)
5. Kaddis, C.S., Loo, J.A.: Native protein MS and ion mobility: large flying proteins with ESI. Anal. Chem. **79**, 1778–1784 (2007)
6. Saucy, D.A., Ude, S., Lenggoro, I.W., Fernandez de la Mora, J.: Mass analysis of water-soluble polymers by mobility measurement of charge-reduced ions generated by electrosprays. Anal. Chem. **76**, 1045–1053 (2004)
7. Müller, R., Laschober, C., Szymanski, W.W., Allmaier, G.: Determination of molecular weight, particle Size, and density of high number generation PAMAM dendrimers using MALDI-TOF-MS and nES-GEMMA. Macromolecules **40**, 5599–5605 (2007)
8. Kemptner, J., Marchetti-Deschmann, M., Siekmann, J., Turecek, P.L., Schwarz, H.P., Allmaier, G.: GEMMA and MALDI-TOF MS of reactive PEGs for pharmaceutical applications. J. Pharm. Biomed. Anal. **52**, 432–437 (2010)
9. Pease III, L.F., Lipin, D.I., Tsai, D.H., Zachariah, M.R., Lua, L.H., Tarlov, M.J., Middelberg, A.P.J.: Quantitative characterization of virus-like particles by asymmetrical flow field flow fractionation, electrospray differential mobility analysis, and transmission electron microscopy. Biotechnol. Bioeng. **102**, 845–855 (2009)
10. Pease III, L.F.: Physical analysis of virus particles using electrospray differential mobility analysis. Trends Biotechnol. **30**, 216–224 (2012)
11. Pease, L.F., Tsai, D.-H., Zangmeister, R.A., Zachariah, M.R., Tarlov, M.J.: Quantifying the surface coverage of conjugate molecules on functionalized nanoparticles. J. Phys. Chem. C **111**, 17155–17157 (2007)
12. Tsai, D.H., DelRio, F.W., Keene, A.M., Tyner, K.M., MacCuspie, R.I., Cho, T.J., Zachariah, M.R., Hackley, V.A.: Adsorption and conformation of serum albumin protein on gold nanoparticles investigated using dimensional measurements and in situ spectroscopic methods. Langmuir **27**, 2464–2477 (2011)
13. Hinterwirth, H., Wiedmer, S.K., Moilanen, M., Lehner, A., Allmaier, G., Waitz, T., Lindner, W., Laemmerhofer, M.: Comparative method evaluation for size and size-distribution analysis of gold nanoparticles. J. Sep. Sci. **36**, 2952–2961 (2013)
14. Fissan, H., Kennedy, M.K., Krinke, T.J., Kruis, F.E.: Nanoparticles from the gas phase as building blocks for electrical devices. J. Nanopart. Res. **5**, 299–310 (2003)
15. Kallinger, P., Weiss, V.U., Lehner, A., Allmaier, G., Szymanski, W.W.: Analysis and handling of bio-nanoparticles and environmental nanoparticles using electrostatic aerosol mobility. Particuology **11**, 14–19 (2013)
16. Havlik, M., Marchetti-Deschmann, M., Friedbacher, G., Winkler, W., Messner, P., Perez-Burgos, L., Tauer, C., Allmaier, G.: Comprehensive size-determination of whole virus vaccine particles using gas-phase electrophoretic mobility macromolecular analyzer, atomic force microscopy, and transmission electron microscopy. Anal. Chem. **87**, 8657–8664 (2015)
17. Loo, J.A., Berhane, B., Kaddis, C.S., Wooding, K.M., Xie, Y., Kaufman, S.L., Chenushevich, I.V.: Electrospray ionization mass spectrometry and ion mobility analysis of the 20S proteasome complex. J. Am. Soc. Mass Spectrom. **16**, 998–1008 (2005)
18. de la Mora, J.F., Ude, S., Thomson, B.A.: The potential of differential mobility analysis coupled to MS for the study of very large singly and multiply charged proteins and protein complexes in the gas phase. Biotechnol. J. **1**, 988–997 (2006)
19. Kaddis, C.S., Lomeli, S.H., Yin, S., Berhane, B., Apostol, M.I., Kickhoefer, V.A., Rome, L.H., Loo, J.A.: Sizing large proteins and protein complexes by electrospray ionization mass spectrometry and ion mobility. J. Am. Soc. Mass Spectrom. **18**, 1206–1216 (2007)

20. Pease 3rd, L.F., Elliott, J.T., Tsai, D.H., Zachariah, M.R., Tarlov, M.J.: Determination of protein aggregation with differential mobility analysis: application to IgG antibody. Biotechnol. Bioeng. **101**, 1214–1222 (2008)
21. Guha, S., Wayment, J.R., Tarlov, M.J., Zachariah, M.R.: Electrospray-differential mobility analysis as an orthogonal tool to size-exclusion chromatography for characterization of protein aggregates. J. Pharm. Sci. **101**, 1985–1994 (2012)
22. Shahidi-Noghabi, S., Van Damme, E.J., Smagghe, G.: Carbohydrate-binding activity of the type-2 ribosome-inactivating protein SNA-I from elderberry (*Sambucus nigra*) is a determining factor for its insecticidal activity. Phytochemistry **69**, 2972–2978 (2008)
23. Schwefel, D., Maierhofer, C., Beck, J.G., Seeberger, S., Diederichs, K., Moller, H.M., Welte, W., Wittmann, V.: Structural basis of multivalent binding to wheat germ agglutinin. J. Am. Chem. Soc. **132**, 8704–8719 (2010)
24. Brewer, C.F., Bhattacharyya, L.: Specificity of concanavalin A binding to asparagine-linked glycopeptides. A nuclear magnetic relaxation dispersion study. J. Biol. Chem. **261**, 7306–7310 (1986)
25. Mandal, D.K., Kishore, N., Brewer, C.F.: Thermodynamics of lectin-carbohydrate interactions. Titration microcalorimetry measurements of the binding of N-linked carbohydrates and ovalbumin to concanavalin A. Biochemistry **33**, 1149–1156 (1994)
26. Giménez, E., Benavente, F., Barbosa, J., Sanz-Nebot, V.: Towards a reliable molecular mass determination of intact glycoproteins by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. **21**, 2555–2563 (2007)
27. Fu, D., van Halbeek, H.: N-glycosylation site mapping of human serotransferrin by serial lectin affinity chromatography, fast atom bombardment-mass spectrometry, and 1H nuclear magnetic resonance spectroscopy. Anal. Biochem. **206**, 53–63 (1992)
28. Sturiale, L., Barone, R., Palmigiano, A., Ndosimao, C.N., Briones, P., Adamowicz, M., Jaeken, J., Garozzo, D.: Multiplexed glycoproteomic analysis of glycosylation disorders by sequential yolk immunoglobulins immunoseparation and MALDI-TOF MS. Proteomics **8**, 3822–3832 (2008)
29. Blanchard, V., Liu, X., Eigel, S., Kaup, M., Rieck, S., Janciauskiene, S., Sandig, V., Marx, U., Walden, P., Tauber, R., Berger, M.: N-glycosylation and biological activity of recombinant human alpha1-antitrypsin expressed in a novel human neuronal cell line. Biotechnol. Bioeng. **108**, 211–2128 (2011)
30. Nakano, M., Kakehi, K., Tsai, M.H., Lee, Y.C.: Detailed structural features of glycan chains derived from alpha1-acid glycoproteins of several different animals: the presence of hypersialylated, O-acetylated sialic acids but not disialyl residues. Glycobiology **14**, 431–441 (2004)
31. Matthews, B.W.: The structure of *E. coli* beta-galactosidase. C. R. Biol. **328**, 549–556 (2005)
32. Signor, L., Boeri Erba, E.: Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometric analysis of intact proteins larger than 100 kDa. J. Vis. Exp. **50635**, (2013)
33. Van Damme, E.J., Barre, A., Rouge, P., Van Leuven, F., Peumans, W.J.: The NeuAc(α-2,6)-Gal/GalNAc-binding lectin from elderberry (*Sambucus nigra*) bark, a type-2 ribosome-inactivating protein with an unusual specificity and structure. Eur. J. Biochem. **235**, 128–137 (1996)
34. Weiss, V.U., Kolivoska, V., Kremser, L., Gas, B., Blaas, D., Kenndler, E.: Virus analysis by electrophoresis on a microfluidic chip. J. Chromatogr. B Anal. Technol. Biomed. Life Sci. **860**, 173–179 (2007)
35. Broekaert, W.F., Nsimba-Lubaki, M., Peeters, B., Peumans, W.J.: A lectin from elder (*Sambucus nigra* L.) bark. Biochem. J. **221**, 163–169 (1984)
36. Engel, N., Weiss, V.U., Wenz, C., Rufer, A., Kratzmeier, M., Gluck, S., Marchetti-Deschmann, M., Allmaier, G.: Challenges of glycoprotein analysis by microchip capillary gel electrophoresis. Electrophoresis **36**, 1754–1758 (2015)
37. Kapellios, E.A., Karamanou, S., Sardis, M.F., Aivaliotis, M., Economou, A., Pergantis, S.A.: Using nanoelectrospray ion mobility spectrometry (GEMMA) to determine the size and relative molecular mass of proteins and protein assemblies: a comparison with MALLS and QELS. Anal. Bioanal. Chem. **399**, 2421–2433 (2011)
38. Weiss, V.U., Kerul, L., Kallinger, P., Szymanski, W.W., Marchetti-Deschmann, M., Allmaier, G.: Liquid phase separation of proteins based on electrophoretic effects in an electrospray setup during sample introduction into a gas-phase electrophoretic mobility molecular analyzer (CE-GEMMA/CE-ES-DMA). Anal. Chim. Acta **841**, 91–98 (2014)
39. Allmaier, G., Laschober, C., Szymanski, W.W.: Nano ES GEMMA and PDMA, new tools for the analysis of nanobioparticles-protein complexes, lipoparticles, and viruses. J. Am. Soc. Mass Spectrom. **19**, 1062–1068 (2008)
40. Allmaier, G., Maisser, A., Laschober, C., Messner, P., Szymanski, W.W.: Parallel differential mobility analysis for electrostatic characterization and manipulation of nanoparticles and viruses. TrAC **30**, 123–132 (2011)
41. Rosser, S., de la Mora, J.F.: Vienna-type DMA of high resolution and high flow rate. Aerosol Sci. Technol. **39**, 1191–1200 (2005)
42. Wiedensohler, A.: An approximation of the bipolar charge distribution for particles in the sub-micron size range. J. Aerosol Sci. **19**, 387–389 (1988)

Supplementary Material

nES GEMMA Analysis of Lectins and their Interactions with Glycoproteins – Separation, Detection, and Sampling of the Noncovalent Biospecific Complexes

[Running Title: nES GEMMA of Lectin-Glycoprotein Complexes]

Nicole Y. Engel, Victor U. Weiss, Martina Marchetti-Deschmann, Günter Allmaier

Institute of Chemical Technologies and Analytics, TU Wien, Vienna, Austria

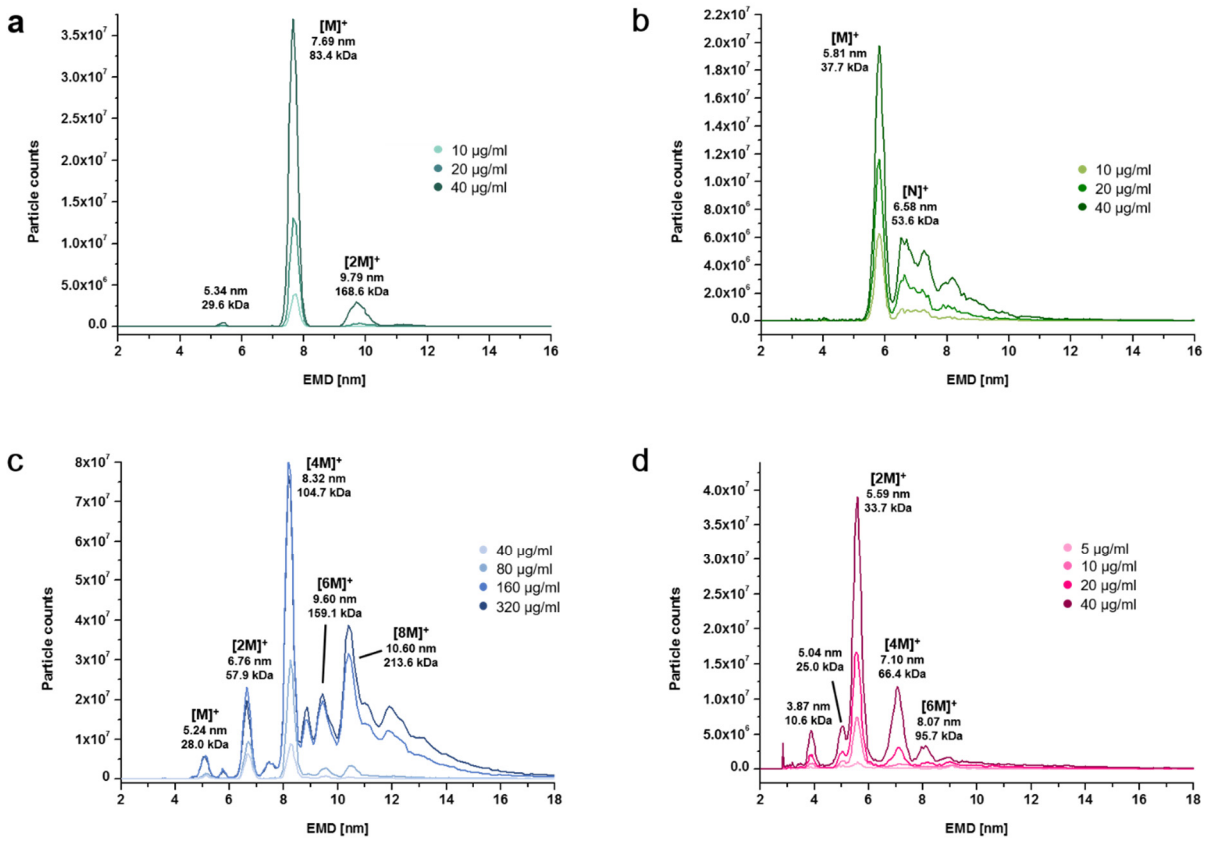
Address reprint requests to:

Prof. Günter Allmaier, Institute of Chemical Technologies and Analytics, TU Wien (Vienna University of Technology), Getreidemarkt 9/164-IAC, A-1060 Vienna, Austria

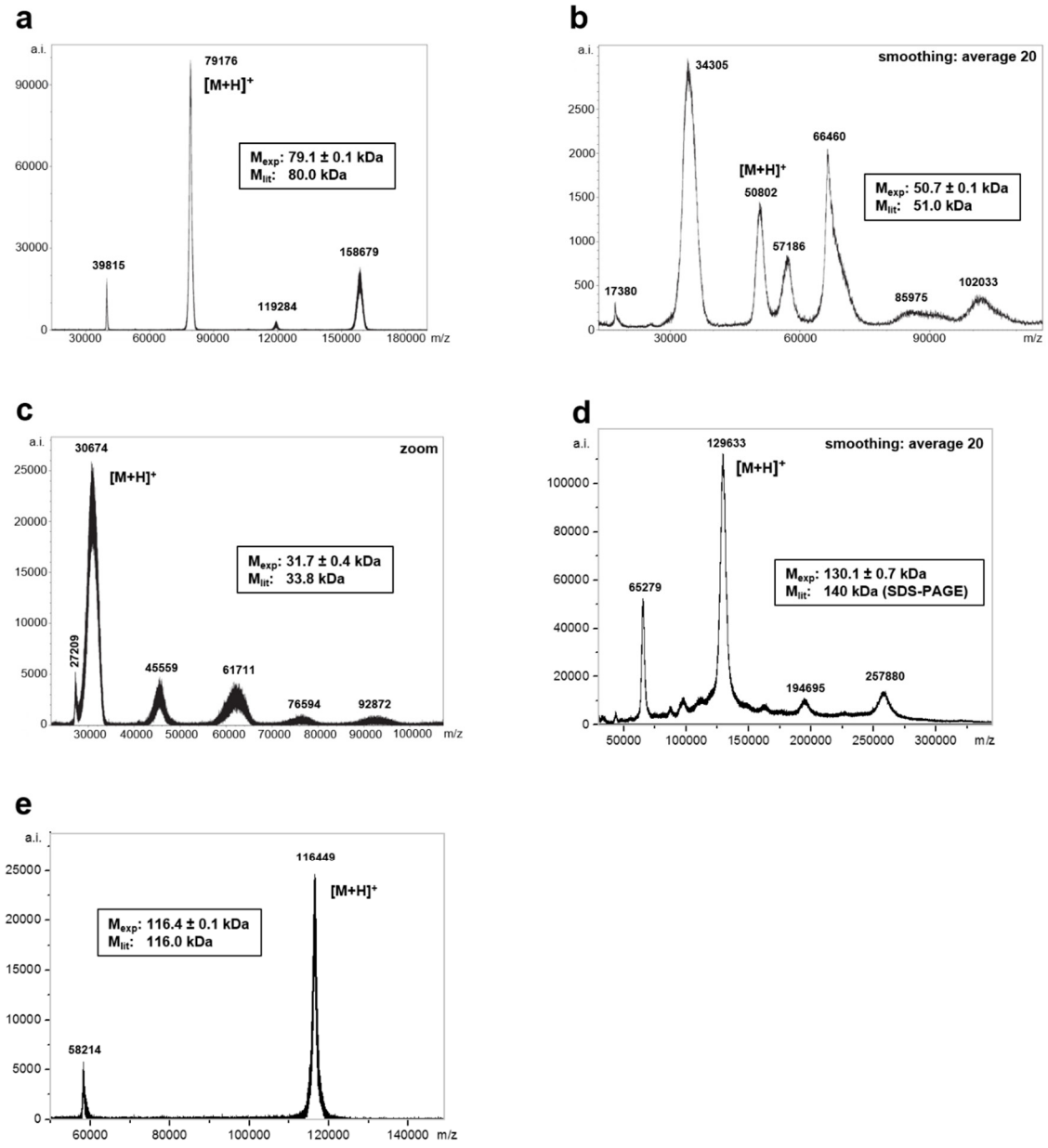
E-mail: guenter.allmaier@tuwien.ac.at

Phone: +43 - 1 - 58801 - 15160

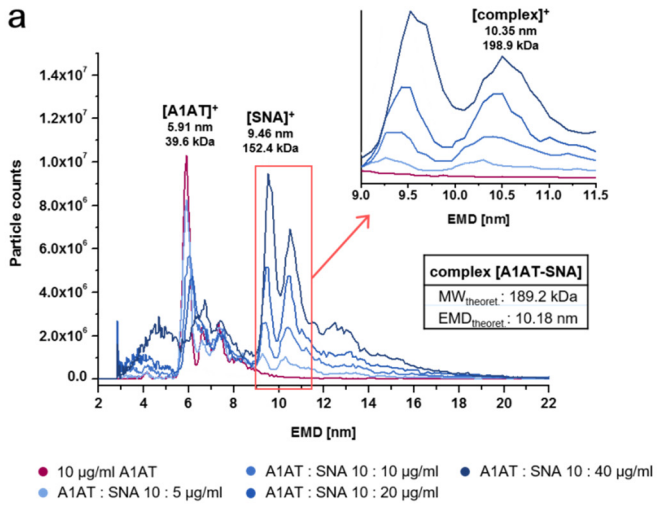
Fax: +43 - 1 - 58801 - 15199

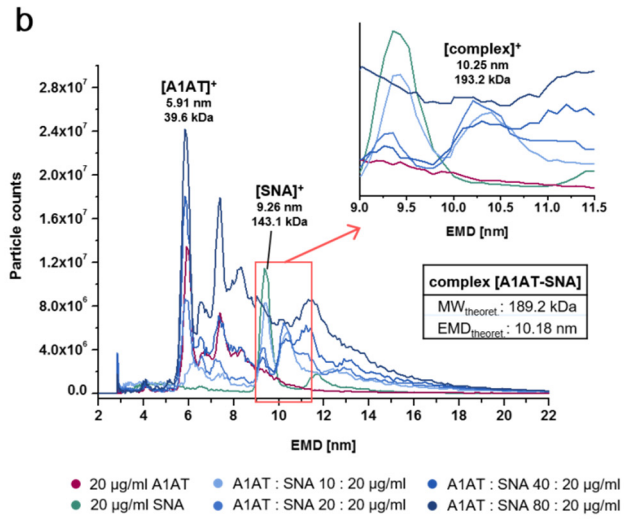


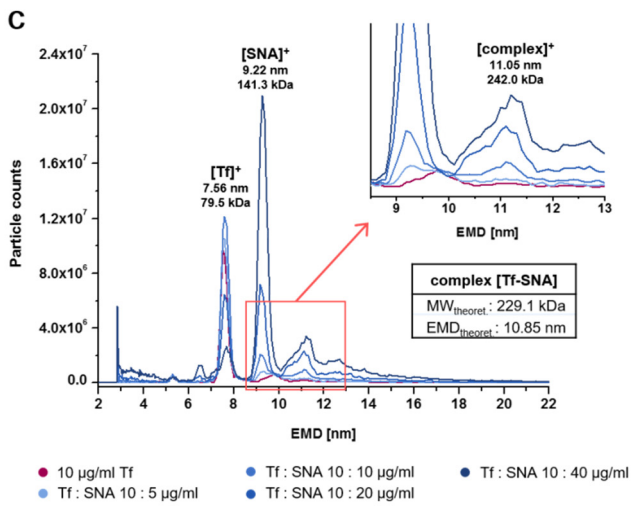
Supplementary Figure 1. nES GEMMA analysis of different concentrations of the glycoproteins Tf (**a**) and A1AT (**b**), as well as the lectins ConA (**c**) and WGA (**d**). $[N]^+$ represents a second constituent of A1AT (**b**).



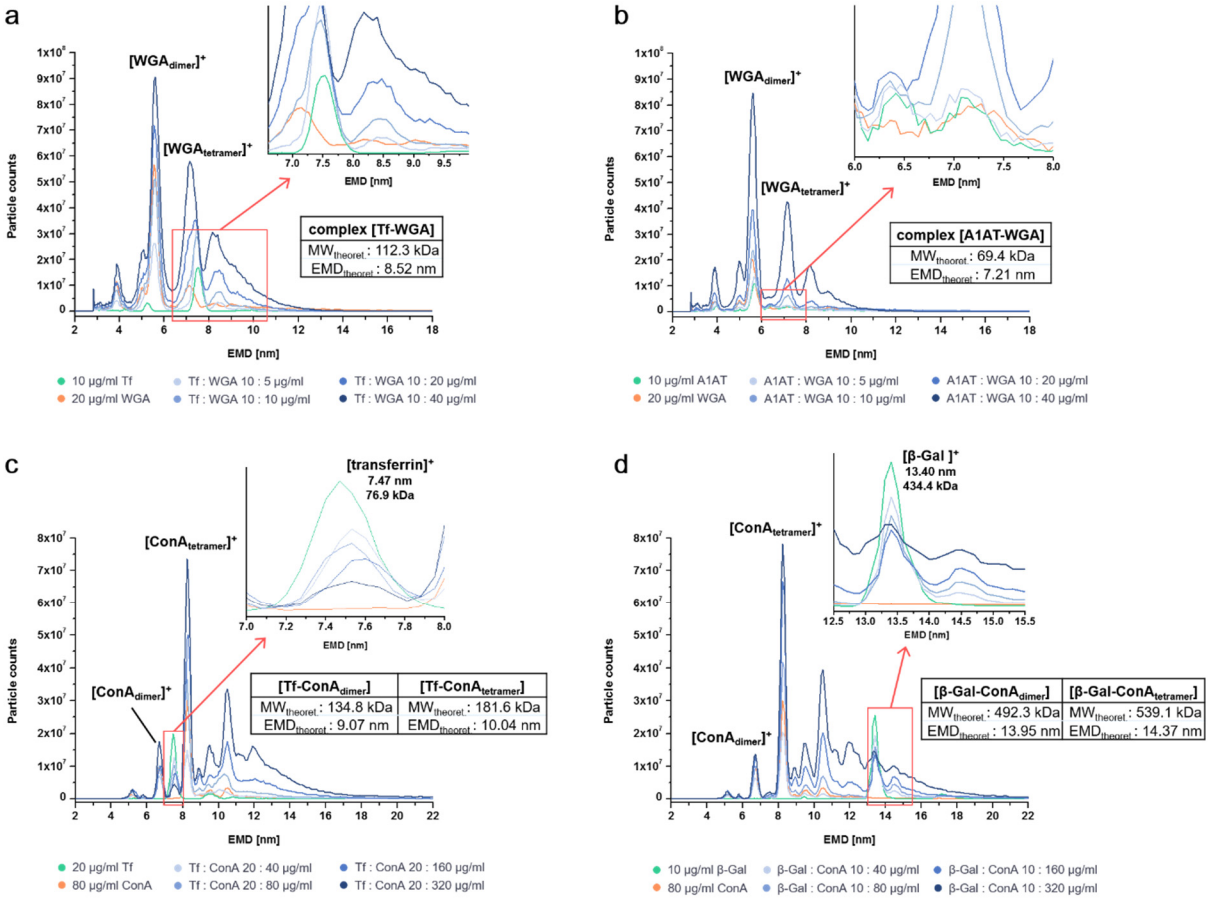
Supplementary Figure 2. Positive ion MALDI mass spectra of 10 pmol Tf (**a**), 10 pmol A1AT (**b**), 10 pmol AGP (**c**), 20 pmol SNA (**d**), and 1.5 pmol β-Gal (**e**) on target.



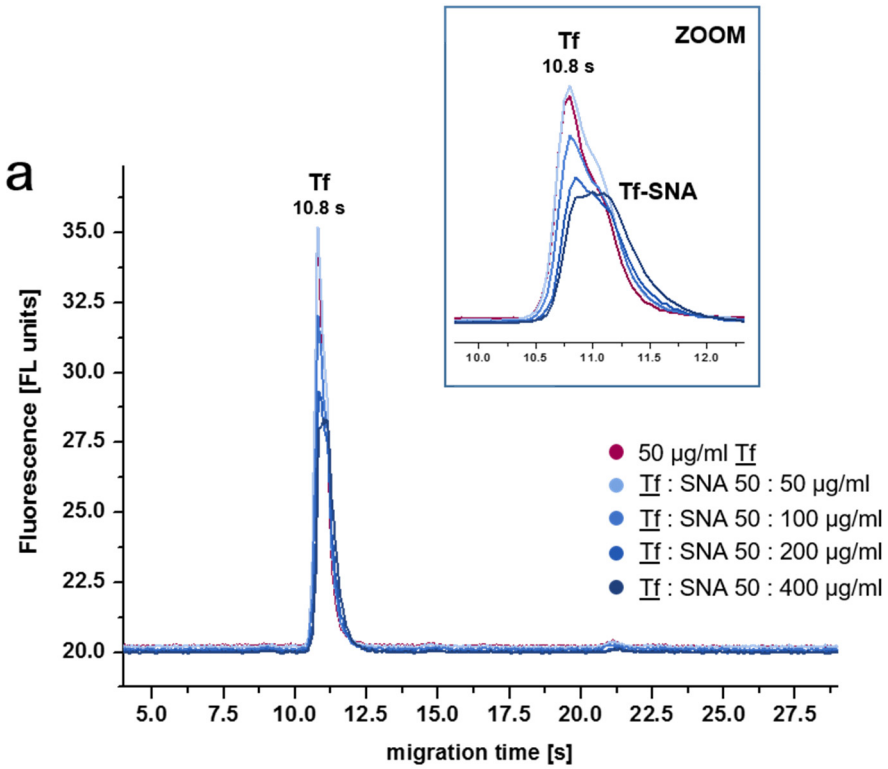


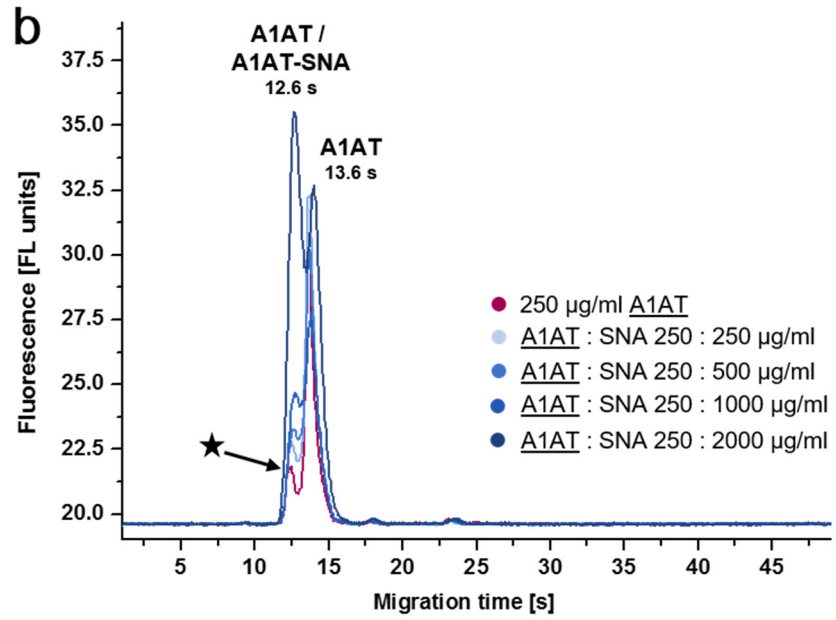


Supplementary Figure 3. nES GEMMA analysis of A1AT (**a**) and Tf (**c**) incubated with different concentrations of SNA, as well as SNA incubated with different concentrations of A1AT (**b**).

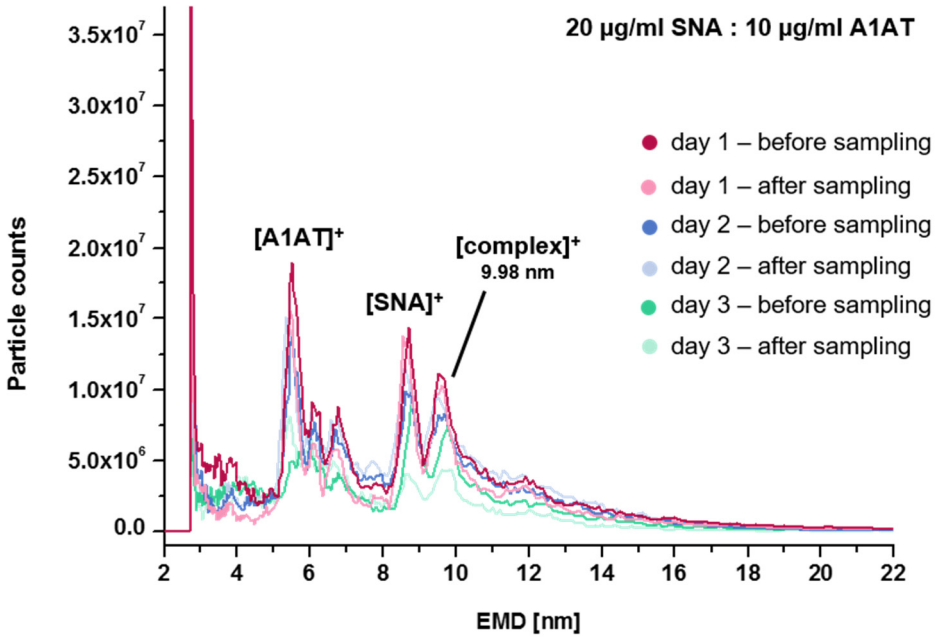


Supplementary Figure 4. nES GEMMA analysis of different concentrations of WGA incubated with the glycoproteins Tf (**a**) and A1AT (**b**) and of ConA incubated with Tf (**c**) and β-Gal as negative control (**d**).





Supplementary Figure 5. CE-on-a-chip interaction analysis of labeled Tf (**a**) and A1AT (**b**) with rising concentrations of unlabeled SNA. Labeled proteins are underlined. An unknown constituent of A1AT is marked with an asterisk (*).



Supplementary Figure 6. Sampling of SNA-A1AT complexes onto NC at 9.96 – 10.05 nm for 36 h on three consecutive days using an aerosol sampler.

Agilent Application Note (**2013**), 5991-3435EN

A comparative study of analytical parameters for proteins with different degrees of glycosylation

Nicole Engel[1], Victor U. Weiss[1], Christian Wenz[2], Andreas Rüfer[2], Martin Kratzmeier[2], Susanne Glück[2], Martina Marchetti-Deschmann[1], and Günter Allmaier[1]

[1]Institute of Chemical Technologies and Analytics, Vienna University of Technology, Vienna, Austria
[2]Agilent Technologies R&D and Marketing GmbH & Co. KG, Waldbronn, Germany

A Comparative Study of Analytical Parameters for Proteins with Different Degrees of Glycosylation

Application Note

Authors

Nicole Engel, Victor U. Weiss, Martina Marchetti-Deschmann, and Günter Allmaier
Research Group Bio- and Polymer Analysis, Institute of Chemical Technology and Analytics, Vienna University of Technology
Vienna, Austria

Christian Wenz, Andreas Rüfer, Martin Kratzmeier, and Susanne Glück
Agilent Technologies R&D and Marketing GmbH & Co. KG
Waldbronn, Germany

Abstract

Glycosylation, either co- or post-translational, is one important form of protein modification conveying protein stability, as well as contributing to folding and recognizability. Consequently, glycoproteins are the focus of basic research, bioengineering, and biotechnology. This Application Note describes the analysis of three biological relevant glycoproteins using the Agilent 2100 Bioanalyzer system. Human transferrin, human antitrypsin, and bovine acid glycoprotein were examined with the Agilent High Sensitivity Protein 250 and Protein 230 assays in regard to molecular weight determination (sizing), dynamic range for quantitation, and sensitivity for the selected analytes. Selected glycoproteins vary in their degree of glycosylation, influencing their behavior during capillary gel electrophoresis and sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE). Both methods were compared with respect to characteristic analytical parameters. In general, microchip capillary gel electrophoresis (MCGE) separated the selected glycoproteins in less than a minute with high sensitivity and high reproducibility. However, the determined molecular weights of all analytes exceeded the theoretical values. Moreover, the occurrence of broadening peaks correlated to the degree of glycosylation.

Agilent Technologies

Introduction

The attachment of sugar moieties to proteins greatly influences their stability, solubility, folding, and most importantly, their bioactivity. Moreover, glycosylation patterns have been found to play a major role in various biological processes concerning immunology, development, and cancer biology[1]. To gain a better understanding of the biological function of a glycoprotein, or to control the integrity of an engineered product, it is of utmost importance to find new bioanalytical methods with high sensitivity and reliability to characterize such analytes in a fast manner.

The Agilent 2100 Bioanalyzer system is capable of rapid on-chip protein analysis with high reproducibility and accuracy regarding quantitation and sizing. It offers two different assays, the Protein 230 (P230) and the High Sensitivity Protein 250 (HSP-250) kits, for the electrophoretic separation of proteins up to 230 and 250 kDa, respectively. Electrophoresis of denatured proteins is carried out in a sieving matrix and buffer containing SDS. Sensitivities of the assays can exceed Coomassie Blue and silver-stained SDS-PAGE analyses[2,3]. Furthermore, both assays exhibit similar levels of reproducibility and accuracy, but differ in regard to labeling procedure and sensitivity. Proteins analyzed with the P230 Kit are not directly labeled but entrapped in SDS micelles containing fluorescence dye. In contrast, in the HSP-250 assay, proteins are covalently labeled with the HSP-250 dye prior to electrophoretic separation showing overall higher sensitivity.

To assess the electrophoretic behavior of glycoproteins with the 2100 Bioanalyzer system, three biological relevant proteins with varying glycosylation moieties were examined with both assays and compared to traditional SDS-PAGE. The characteristics of the selected glycoproteins, human transferrin, human antitrypsin, and bovine acid glycoprotein, are summarized in Table 1. All samples were analyzed with respect to reproducibility and accuracy of sizing and quantitation, sensitivity, peak pattern, and peak shape.

Table 1. Summary of the characteristics of human transferrin, human antitrypsin, and bovine acid glycoprotein.

Glycoprotein	Amino acids	N-glycosylation sites	MW_{lit} (MS) (kDA)	MW_{lit} (SDS-PAGE) (kDA)	Approximate glycan content
Transferrin[4,5]	697	Asn^{413}, Asn^{611}	80	80	6 %
Antitrypsin[6,7]	394	Asn^{46}, Asn^{83}, Asn^{247}	51	52	13 %
Acid glycoprotein[8]	184	Asn^{16}, Asn^{39}, Asn^{76}, Asn^{86}, Asn^{118}	33.8	42	37 %

Experimental

Materials

Human serum transferrin (≥ 98 %), bovine acid glycoprotein (99 %), human antitrypsin (salt free, lyophilized powder), dithiothreitol (DTT, BioUltra), silver nitrate and sodium carbonate (both analytical grade), as well as formaldehyde solution (35 % wt., for molecular biology) were purchased from Sigma-Aldrich (Taufkirchen, Germany). Sodium thiosulfate pentahydrate, ethanol, and acetic acid (all analytical grade) were obtained from Merck (Darmstadt, Germany). NuPAGE 4–12 % *Bis-Tris* gels, 4x lithium dodecyl sulfate (LDS) sample buffer (106 mM *Tris* HCl, 141 mM *Tris* Base, 2 % LDS, 10 % Glycerol, 0.51 mM EDTA, 0.22 mM Serva Blue G250, 0.175 mM Phenol Red, pH 8.5), 20x MES SDS running buffer (50 mM MES, 50 mM *Tris* Base, 0.1 % SDS, 1 mM EDTA, pH 7.3), and BenchMark Protein Ladder were purchased from Invitrogen (Darmstadt, Germany). The Agilent 2100 Bioanalyzer system, P230 and HSP-250 kits were obtained from Agilent Technologies (Waldbronn, Germany).

SDS-PAGE

Samples in 1x LDS sample buffer and 50 mM DTT were run on NuPAGE 4–12 % *Bis-Tris* gels in 1x MES SDS running buffer at 120 V (const.) and 60 mA (max.). For molecular weight determination, BenchMark Protein Ladder was applied. Protein bands were seen by silver staining[9]. The gel was washed with 50 % ethanol and 5 % acetic acid for 20 minutes, with 50 % ethanol for 10 minutes, and three times with water for 20 minutes each. Afterwards, the gel was incubated in 0.2 % sodium thiosulfate pentahydrate for 1 minute and washed twice with water for 1 minute each. Subsequently, the gel was incubated with 1 % silver nitrate at 4 °C for 20 minutes and again washed twice with water for 1 minute each. The final incubation was in 2 % sodium carbonate and 0.04 % formaldehyde solution until protein bands were visible. Staining was stopped with 5 % acetic acid.

Protein labeling for HSP-250 analysis

All proteins and the HSP-250 ladder were labeled according to the description in the Agilent High Sensitivity Protein 250 Kit Guide[10]. Briefly, 10 µL of protein solution was mixed with 1 µL HSP-250 labeling dye and incubated on ice for 30 minutes. The labeling reaction was stopped by adding 1 µL of ethanolamine and incubating on ice for 10 minutes. For reproducibility and sensitivity experiments, 700 µg/mL glycoprotein solutions were labeled and diluted to the required concentration for chip analysis.

On-chip analysis

The chip-based glycoprotein analysis was performed on the Agilent 2100 Bioanalyzer system either with the P230 or the HSP-250 assay. Each chip was prepared according to the respective kit guide[10,11]. Briefly, samples were diluted 1:200 with water prior to analysis with the HSP-250 assay. Aliquots of 4 µL were incubated with 2 µL of sample buffer in the presence of DTT at 95 °C for 5 minutes. Samples analyzed with the HSP-250 assay were directly applied

to the chip, whereas 84 µL water were added to the samples used with the P230 assay before applying 6 µL of the dilution to the chip.

Results and Discussion

All three glycoproteins were analyzed with the 2100 Bioanalyzer system using the P230 and the HSP-250 kit, as well as with SDS-PAGE and subsequent silver staining. Glycoprotein sizing on a MCGE system, as with the 2100 Bioanalyzer solution, faces the same pre-analytical problems as traditional slab gel electrophoresis. Hydrophilic and bulky carbohydrate moieties on the amino acid backbone lead to a hampered interaction with detergents in comparison to unmodified proteins. In MCGE and SDS-PAGE, it can be observed that glycoproteins migrate slower in gel electrophoresis and, therefore, are detected at higher apparent molecular weight (MW)[12].

Figure 1 displays SDS-PAGE gels and gel-like images of electrophoretic separations of transferrin, antitrypsin, and acid glycoprotein with the P230 and the HSP-250 kit, respectively. Having two N-glycosylation sites, transferrin, with a MW of 80 kDa, represents the protein with the lowest glycan content (approximately 6 % of total MW) of the three investigated analytes (Table 1). Whereas a single band with the expected MW of 80 kDa is visible by SDS-PAGE, MCGE gives an average MW (MWexp) of ~90 kDa (89.3 ± 1.3 kDa for P230 and 90.4 ± 1.4 kDa for HSP-250) for both assays equaling a MW overdetermination of about 12.5 % (Figure 1A).

The average MW was calculated for both assays from over 15 measurements. For MCGE, the relevant signals are sharp and distinct (small peak width), as seen in the exemplary electropherograms presented in Figures 2A and 2C. The peak shape can be compared to nonglycosylated proteins (for example, the protein ladder which is always applied on the same protein chip as the samples).

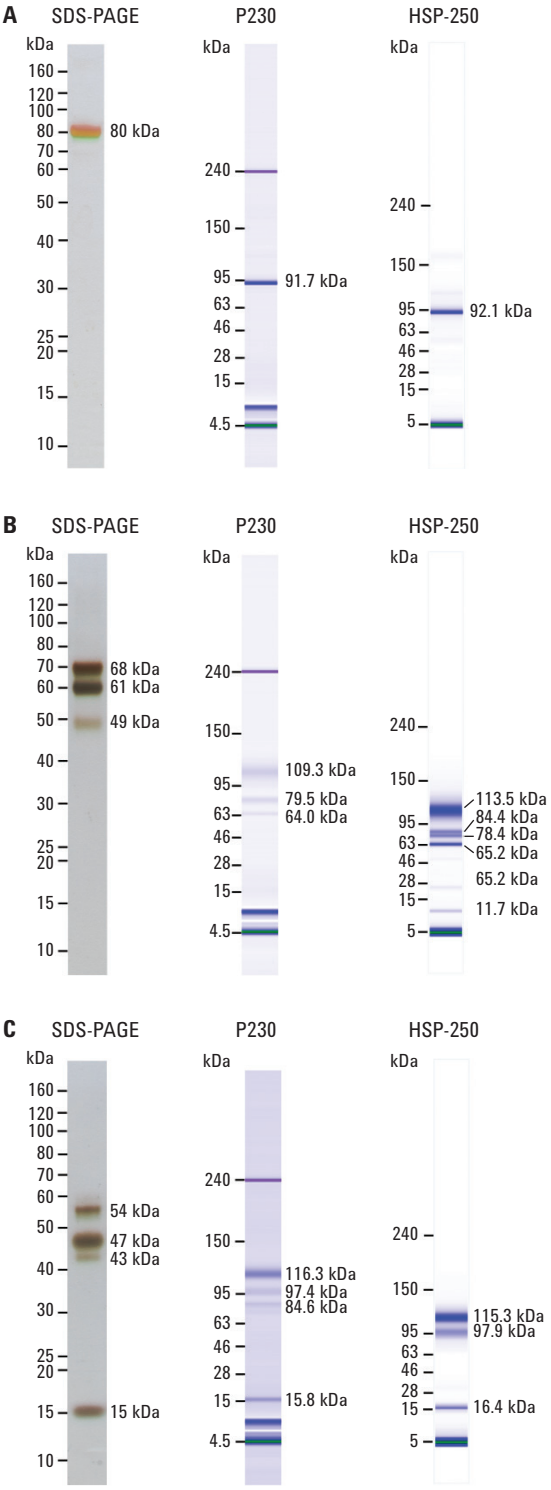


Figure 1. Comparison of the electrophoretic separation of (A) human transferrin, (B) human antitrypsin, and (C) bovine acid glycoprotein with SDS-PAGE, P230, and HSP-250 assays. For SDS-PAGE analysis, 250 ng of transferrin (A), 695 ng of antitrypsin (B), and 348 ng of acid glycoprotein (C) were applied. For the Agilent 2100 Bioanalyzer measurements, 550 µg/mL of each glycoprotein were applied.

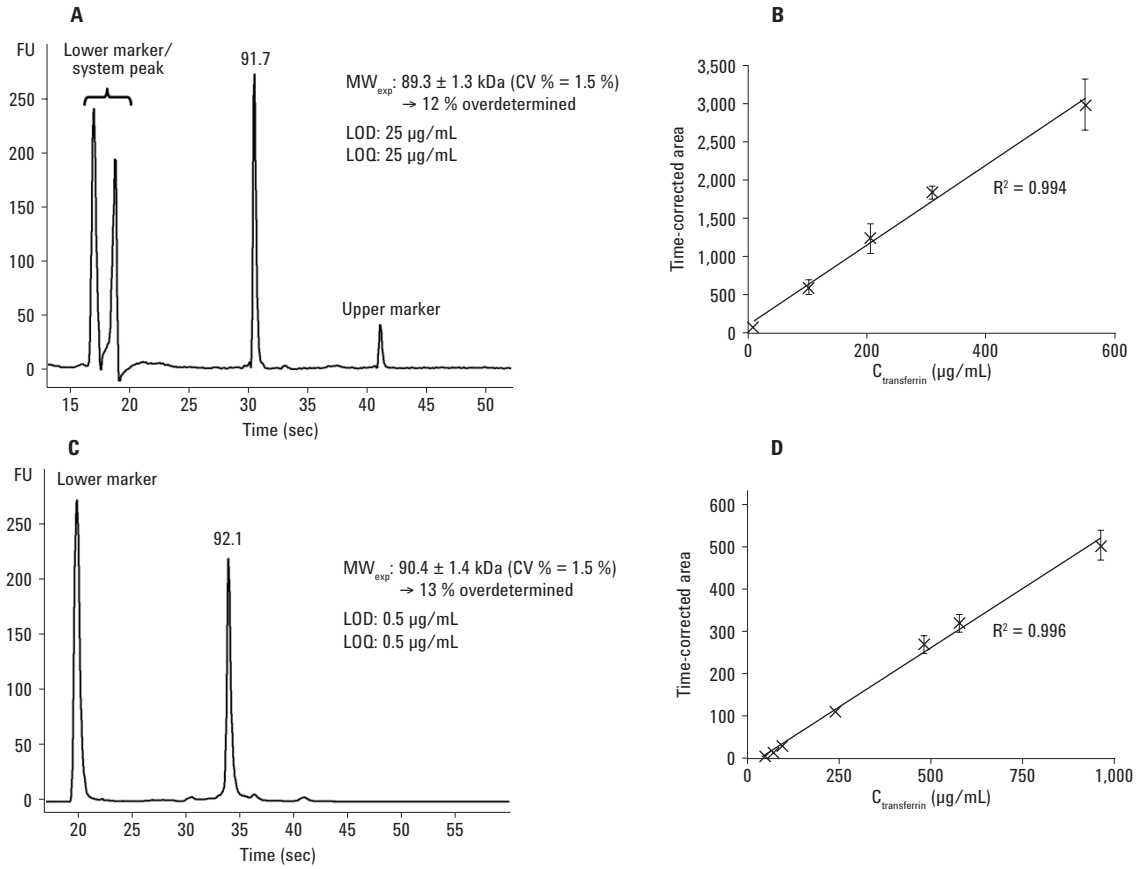


Figure 2. Gel electrophoretic analysis of transferrin on the Agilent 2100 Bioanalyzer system. 550 µg/mL transferrin was separated using (A) the P230 and (C) the HSP-250 assays (exemplary results depicted). The correlation coefficient (R^2) was determined to be (B) 0.994 for the P230 and (D) 0.996 for the HSP-250 assays.

Antitrypsin, which exhibits three N-glycosylation sites, has an expected MW of 51 kDa and an approximate glycan content of 13 % of the overall MW (Table 1). The SDS-PAGE gel shows three distinct bands with MWs of 49, 61, and 68 kDa (Figure 1B), respectively. These three bands can also be detected with the P230 assay, yet the determined average MWs of 63.4 ± 1.0, 78.8 ± 1.0, and 108.1 ± 1.4 kDa exceeded the expected values by over 100 % in the most abundant component (Figure 3A). A further increase in MW was observed for the HSP-250 run (65.0 ± 0.8, 84.3 ± 1.2, and 112.0 ± 1.2 kDa), which can be explained by the altered behavior of labeled glycoproteins compared to nonlabeled ones. The highly sensitive assay allowed for the detection of a fourth component with a MW of 78.0 ± 1.3 kDa not visible by SDS-PAGE (Figure 3C). In comparison to transferrin, the antitrypsin signals of the 2100 Bioanalyzer measurements were broader without causing greater deviations in sizing (approximately 1.1–1.6 % for each signal).

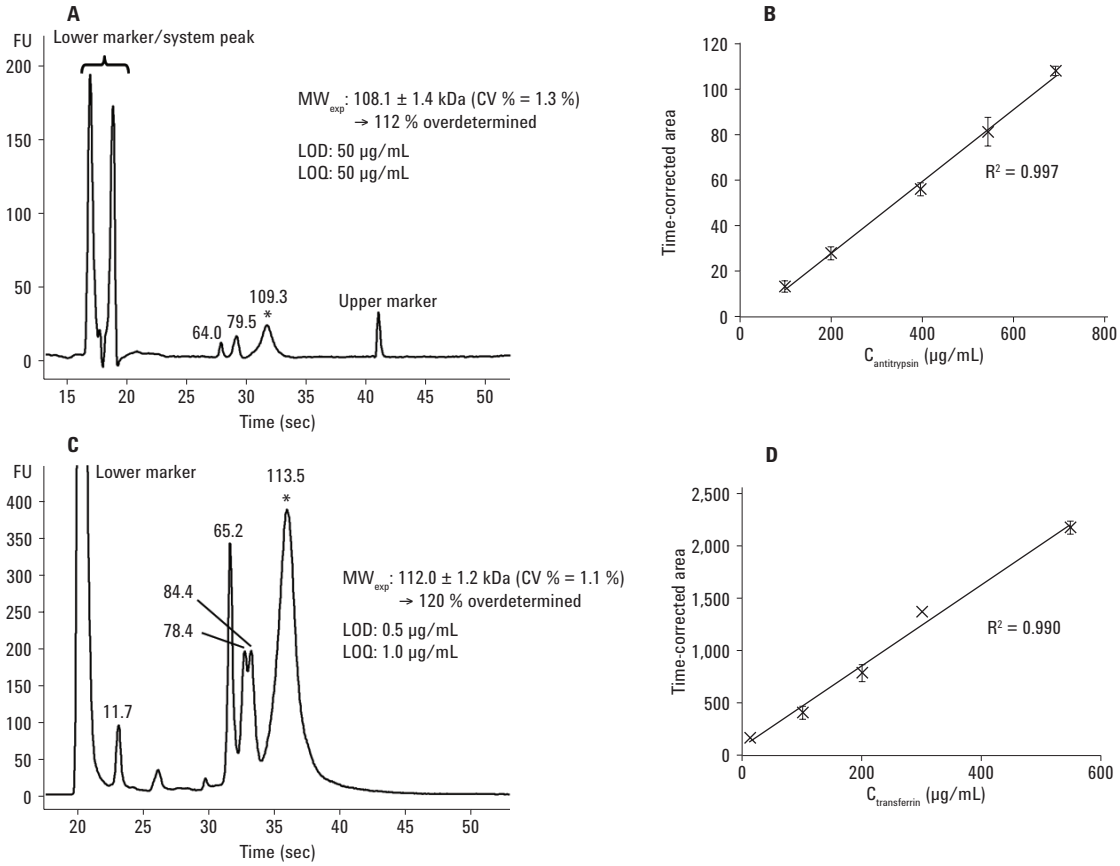


Figure 3. Gel electrophoretic analysis of antitrypsin with the Agilent 2100 Bioanalyzer system. 550 µg/mL antitrypsin was separated using (A) the P230 and (C) the HSP-250 assays (exemplary results depicted). The * marks the most abundant and evaluated signal. The correlation coefficient (R^2) was determined to be (B) 0.997 for the P230 and (D) 0.990 for the HSP-250 assays.

The third analyte, acid glycoprotein, has a MW of 33.8 kDa and exhibits five N-glycosylation sites leading to a glycan content of approximately 37 % (Table 1). Due to this high percentage of attached carbohydrates, an even higher deviation of the measured MWs from literature was expected, as well as an increase in peak heterogeneity compared to the two previously investigated glycoproteins. Corresponding SDS-PAGE and 2100 Bioanalyzer results are presented in Figure 1C, confirming these anticipations. Four protein bands with MWs of 15, 43, 47, and 54 kDa could be seen on the SDS-PAGE gel. The 47 kDa compound was the most abundant one. The 2100 Bioanalyzer investigations using the P230 assay also identified four compounds, but with much higher MWs (15.6 ± 0.2, 84.7 ± 0.7, 98.5 ± 1.2, and 116.3 ± 1.2 kDa, respectively). The largest one being the most abundant. In contrast, the HSP-250 assay only separated three compounds exhibiting similar MWs (15.9 ± 0.6, 95.8 ± 1.8, and 113.6 ± 1.3 kDa) when compared to the P230 kit (Figure 4C).

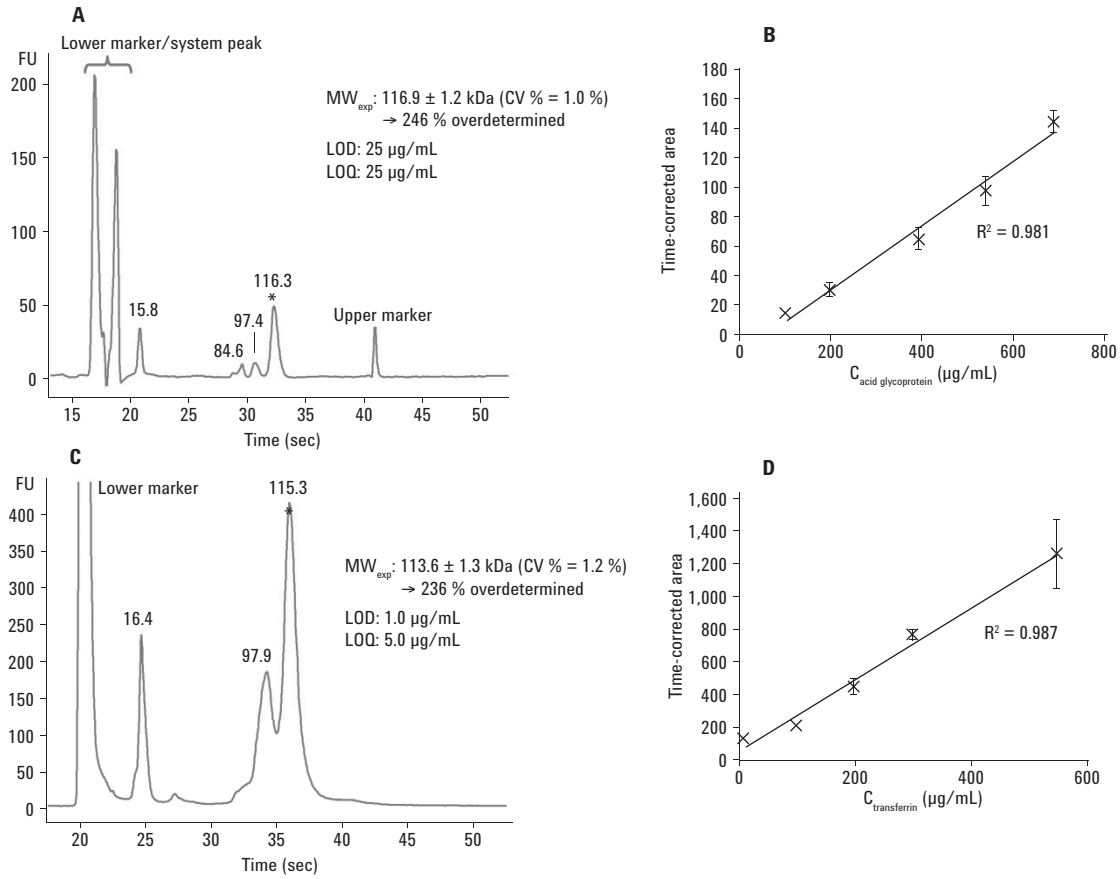


Figure 4. Gel electrophoretic analysis of acid glycoprotein with the Agilent 2100 Bioanalyzer system. 550 µg/mL acid glycoprotein was separated using (A) the P230 and (C) the HSP-250 assays. (exemplary results depicted). The * marks the most abundant and evaluated signal. The correlation coefficient (R^2) was determined to be (B) 0.981 for the P230 and (D) 0.987 for the HSP-250 assays.

To evaluate the sizing reproducibility, sample triplets were measured for each analyte either on three consecutive (inter-chip) or on one 2100 Bioanalyzer chip (intra-chip). Both assays showed low deviations of obtained MW values (less than 1.0 % for intra- and less than 1.7 % for inter-chip reproducibility). Such values lie perfectly within the specifications of both kits with a determined sizing reproducibility of 3 % for measurements with bovine serum albumin[10,11].

The quantitation accuracy and reproducibility were determined for all glycoproteins, and both assays on the 2100 Bioanalyzer system, respectively. Triplets of several different concentrations of each glycoprotein were measured and plotted against corresponding time corrected areas. Figures 2B and 2D display linear regressions of transferrin analyzed with the P230 and the HSP-250 assays. In both cases, good linearity could be determined with correlation coefficients (R^2) of over 0.99 with average deviations of approximately 10 %. The latter was slightly lower for the P230 assay. Antitrypsin also showed good correlation with linear regressions of R^2 being 0.99 or more as demonstrated (Figures 3B and 3D). Deviations of time corrected areas were, on average, approximately 8 % for the P230 and approximately 10 % for the HSP-250 assay, and were always smaller than 17 % for single measurements. In comparison to transferrin and antitrypsin, the correlation coefficients of the linear regressions of acid glycoprotein were slightly lower with values of 0.981 and 0.987 for the P230 and HSP-250 assays, respectively (Figures 4B and 4D). Furthermore, the related deviations were slightly higher for both assays, but always lower than 19 %. Altogether, all glycoproteins could be quantified within the specifications of both assays, which define the quantitation reproducibility with a CV % lower than 20 % as determined for bovine serum albumin[10,11].

Besides the quantitation accuracy and reproducibility, the sensitivity of both assays was investigated for all three analytes. The limit of detection (LOD) defines the concentration with a signal-to-noise ratio (S/N) over 2 and the limit of quantification (LOQ) has an S/N over 3. The specific values for each glycoprotein are presented in Figures 2A/C, 3A/C, and 4A/C. The LOD and LOQ rose with increasing degree of glycosylation in case of the HSP-250 assay. For the P230 assay, antitrypsin showed highest values when compared to transferrin and acid glycoprotein.

Conclusion

Three glycoproteins with varying glycosylation degrees were analyzed on the Agilent 2100 Bioanalyzer system using the P230 and the HSP-250 assays and compared to results from standard SDS-PAGE experiments (Table 2). In general, MWs determined with the 2100 Bioanalyzer system exceeded SDS-PAGE values depending on the degree of glycosylation, as N-glycans interfered with detergent attachment resulting in flawed values of seemingly higher MWs. Additionally, the width of respective peaks broadened in accordance to the number (that is, heterogeneity) of glycan moieties. Consequently, sizing of intact glycoproteins with gel based electrophoretic separation can be difficult. In contrast, quantitation experiments with glycoproteins on the 2100 Bioanalyzer system showed reproducible and accurate results with increasing deviations concomitant to increasing degrees of glycosylation. However, reproducibility of sizing results still lay within the specifications of both assays. Also, the LOD and LOQ tended to show a correlation with the amount of protein carbohydrate modifications.

Table 2. Summary of results for the glycoproteins human transferrin, human antitrypsin and bovine acid glycoprotein for SDS-PAGE, the P230, and HSP-250 assays.

Glyco-protein	Approx. glycan content	MW_{lit} (MS) in kDa	MW_{exp} (SDS-PAGE) in kDa	MW_{exp} (P230) in kDa	MW_{exp} (HSP-250) in kDa	MW_{exp} (HSP-250) in kDa	Average CV % (P230)*	Average CV % (HSP-250)*	LOD/LOQ (P230) in µg/mL	LOD/LOQ (HSP-250) in µg/mL
Transferrin	6 %	80	80	89.3 ± 1.3	90.4 ± 1.4	90.4 ± 1.4	9	10	25/25	0.5/0.5
Antitrypsin	13 %	51	49, 61, 68	63.4 ± 1.0 78.8 ± 1.0 108.1 ± 1.4	65.0 ± 0.8 78.0 ± 1.3 84.3 ± 1.2 112.0 ± 1.2	65.0 ± 0.8 78.0 ± 1.3 84.3 ± 1.2 112.0 ± 1.2	8	10	50/50	0.5/1
Acid glycol-protein[8]	37 %	33.8	15, 43, 47, 54	15.6 ± 0.2 84.7 ± 0.7 98.5 ± 1.2 116.3 ± 1.2	15.9 ± 0.6 95.8 ± 1.8 113.6 ± 1.3	15.9 ± 0.6 95.8 ± 1.8 113.6 ± 1.3	8	11	25/25	1/5

* Average CV %: average deviation of quantitation concerning the evaluated signal.

References

1. A. Varki, *Glycobiology*, 3, 97 (**1993**).
2. L. Bousse, *et al., Anal. Chem*., 73, 1207 (**2001**).
3. "Performance characteristics of the High Sensitivity Protein 250 assay for the Agilent 2100 bioanalyzer" Agilent Technical Note, publication number 5989-8940EN.
4. D. Fu, *et al., Anal. Biochem*., 206, 53 (**1992**).
5. E. Giménez, *et al., Rapid Commun. Mass Spectrom*., 21, 2555 (**2007**).
6. V. Blanchard, *et al., Biotechnol. Bioeng*. 2011, 108, (**2118**).
7. L. Sturiale, *et al., Proteomics*, 8, 3822 (**2008**).
8. M. Nakano, *et al., Glycobiology*, 14, 431 (**2004**).
9. A. Shevchenko, *et al., Anal. Chem.*, 68, 850 (**1996**).
10. Agilent High Sensitivity Protein 250 Kit Guide, Agilent Technologies Manual, publication number G2938-90310.
11. Agilent Protein 230 Kit Guide, Agilent Technologies Manual, publication number G2938-90054.
12. "Glycoprotein sizing on the Agilent 2100 bioanalyzer" Agilent Technologies Application Note, publication number 5989-0332EN.

www.agilent.com/genomics/bioanalyzer

Published in the USA, December 1, 2013
5991-3435EN

Agilent Technologies

TSI Inc Application Note (submitted **2016**)

MacroIMS analysis of a monoclonal antibody – influences of scan time, electrolyte pH, and temperature treatment of sample on obtained signals

Nicole Y. Engel, Victor U. Weiss, Nicole Puffler, Martina Marchetti-Deschmann, Günter Allmaier

Institute of Chemical Technologies and Analytics, TU Wien, Vienna, Austria

UNDERSTANDING, ACCELERATED

Application Note

MacroIMS Analysis of a Recombinant Monoclonal Antibody – Influences of Scan Time, Electrolyte pH, and Temperature Treatment of Sample on Obtained Signals

Nicole Y. Engel, Victor U. Weiss, Nicole Puffler and Günter Allmaier

Institute of Chemical Technology and Analytics, TU Wien, Vienna, Austria

Introduction

Monoclonal antibodies have become major therapeutic and diagnostic agents applied for the treatment and detection of many diseases like cancer, inflammation or infectious diseases as well as for biomedical research and analytics. Thus, it is of importance to find highly sensitive methods to monitor the antibodies' production, purity, and the formation of aggregates or degradation products. Due to working at atmospheric pressure and under relatively mild conditions, the MacroIMS system with the Advanced Electrospray (Model 3482) device based on the soft X-ray charge conditioning concept [1] as successor of nES GEMMA allows the analysis of macromolecules in a native-like environment keeping non-covalent interactions intact [2].

Studying a 148 kDa recombinant monoclonal antibody, the influence of scan times, pH, and temperature on the measurements was evaluated. Special interest was focused on changes in electrophoretic mobility diameter (EMD) values, peak form, and signal intensity.

Sample Preparation and Analytical Conditions

376 nanomole recombinant monoclonal antibody (rmAB) of therapeutic grade dissolved in 5 mM histidine and 60 mM trehalose buffer (pH 6.0) was used. Sample concentration and buffer exchange to 40 mM ammonium acetate (pH 8.0 or 5.0) was carried out by means of 10 kDa centrifugal filters (VWR, Vienna, Austria) and dilutions ranging from 4.7 to 7527.0 nM were prepared.

Samples were delivered via a syringe pump (NE-1000, New Era Pump Systems, Farmingdale, NY, USA) to the MacroIMS system (Model 3982). nES sheath gas (filtered, dried air and CO_2) was set to 1.37 Lpm and voltages were adjusted for stable cone jet mode (1.39-1.45 kV). DMA laminar flow was set to 30 liters per minute with scan times of 11 or 56 s (3 s retrace and 1 s lag time each in order to set the voltage to initial conditions). Scans from 2 to 45 nm EMD were acquired during 20 min of continuous sampling. All data result from triplicate measurements.

Results and Discussion

rmAB dilutions were analyzed with different measurement scan times, 11 and 56 s at two different pHs, 5.0 and 8.0. The shorter scan time of 11 s was chosen for fast analysis times needed for potential *on-line* coupling with high performance liquid chromatography systems (HPLC). 56 s was chosen for optimal resolution. Results were compared in regard to EMD values, peak form, and peak height or area. Limit of detection (LOD) and limit of quantitation (LOQ) were determined for rmAB analysis. Moreover, the influence of temperature treatment of this rmAB sample prior to analysis was studied.

Influence of scan time and pH on reproducibility, EMD, and peak width

Inter-day and intra-day variability was checked for measurements with 56 s scan time. Good repeatability for (i) EMDs, (ii) peak widths, and (iii) peak area/heights were observed. The latter parameter, namely intra-day variation was below 6 % but on the other hand the inter-day variation increased to 28 % (standard deviations (SDs)). Reproducible measurement set-ups applying well-defined parameters such as the same capillary with excellent tip geometry, and the formation of a stable cone-jet mode are of importance for achieving these results. SDs can be improved further, i.e. reduced SD values, by replacing the syringe pump for direct infusion with a flow injection analysis (FIA) system or HPLC system.

Directly comparing different DMA scan times at different pHs shows the influence (**Figure 1A**). Switching from 56 s to 11 s drastically changes the signal intensity, peak width, and EMD. EMDs were detected at significantly higher values at lower scan speed (see **Table 1**). However, a mere variation of the pH at 56 s scan time did not affect the EMDs. In all cases, the determined EMD stayed very reproducible for each condition with SDs below 1.3 %.

The lower scan time of 11 s also caused peak broadening of about 30 %. In contrast, the corresponding FWHM values for each scan time stayed quite constant over the concentration ranges with a slight increase towards lower concentrations. Therefore, the SDs of the mean values were higher than for the EMDs (below 8 % for the monomer rmAB signals and below 15 % for the dimers and trimers of rmAB).

Influence of scan time and pH on signal intensity and LOD/LOQ

As expected the most significant influence for different scan times was observed for peak height and area as already shown in **Figure 1A**. This was studied further in more detail. **Figure 1B** displays a dilution experiment for the rmAB covering a concentration range from 4.7 to 7527.0 nM (**Fig. 1B**, pH 8.0, 56 s). The corresponding peak areas for the monomeric, dimeric, and trimeric rmAB signals were plotted against the rmAB concentration. As expected, signal intensity increased with higher concentrations. However, the calibration was not linear for the full concentration range. The upper limit of the linear range for the antibody monomer was at approx. 1000 nM at 56 s and 3000 nM at 11 s scan speed (**Fig. 1C**). The reason for that is the increased tendency to form gas-phase multimers during the electrospray process [3]. The measured dimers (**Fig. 1D**) and trimers (data not shown), however, showed increased linearity, yet of lesser sensitivity (slope of linear regression). Again much lower signal intensity values were observed for 11 s scan, on average about 26 % lower than for 56 s scan duration. Switching from pH 8.0 to pH 5.0 at 235.3 nM showed slightly elevated peak areas at the same scan duration.

In order to define LOD and LOQ of the analysis, pure background electrolyte was analyzed and its mean value as well as the SD in the range of 9 to 11 nm calculated. LOD was then determined as the mean value plus 3 times SD, LOQ as the mean value plus 10 times SD, which resulted in a LOD and LOQ of 5 and 10 nM, respectively. These values were in good accordance with the measurements of 4.7 and 18.8 nM having a signal to noise ratio over 5 and 20, respectively.

Influence of temperature on the sample

Figure 2 displays the effects of temperature treatment on this particular rmAB sample. Incubation of the sample for 15 min at temperatures from RT up to 70 °C did not seem to influence the stability of this recombinant monoclonal antibody. Above this temperature, i.e. at 80 °C, and the selected time window a drastic decrease was observed for signal intensities. No soluble degradation or aggregation products were detected, leading to the conclusion that sample loss occurs because of sedimentation and adsorption at the inner wall of the reaction tube during incubation.

Conclusions

The analysis of a formulated recombinant monoclonal antibody with the MacroIMS system incorporating the Advanced Electrospray (Model 3482) device based on the soft X-ray charge conditioner exhibited clearly the suitability of the device for such biotechnology applications. Furthermore it was shown that different scan times have a major impact on the quality of the measurements. Therefore, it is important to choose this parameter carefully to gain reproducible and accurate results for quantitative analysis particular when switching from infusion mode to FIA or HPLC mode. A reduction of scan time led to decreased signal intensities with higher FWHM and EMD values,

but short scan times are a necessity to detect peaks eluting from a HPLC system. Longer scan times, on the other hand, can increase resolving power in terms of sizing and sensitivity during an experiment and are therefore favorable measurement parameters for direct infusion. In case of *on-line* coupling with HPLC a balance between optimal resolving power of the DMA and scan time of the DMA (i.e. fitting to the peak width of the chromatographic peak) has to be found.

Next to that, a variation in pH and temperature of the rmAB solution to simulate forced stress conditions was performed and monitored by the MacroIMS system. pH did not affect the here applied antibody whereas a temperature increase above 70 °C for a duration above 10 min drastically reduced the measured the analyte concentration of the solution. Overall, the MacroIMS system demonstrated a high reproducibility for all tested conditions and proved to be a suitable tool for recombinant monoclonal antibody analysis and monitoring of production processes.

References

[1] http://www.tsi.com/macroims-macroion-mobility-spectrometer-3982/

[2] G. Bacher et al., *J. Mass Spectrom.* **2001**, *36*, 1038-1052.

[3] S. L. Kaufman, *J. Aerosol Sci.* **1998**, *29*, 537-552.

Table 1: Influence of scan time and pH on the determined EMDs and their SDs related to the analyzed recombinant monoclonal antibody.

Tested condition: scan time / pH	**$EMD_{monomer}$ [nm]**	**EMD_{dimer} [nm]**	**EMD_{trimer} [nm]**
11 s / pH 8.0	9.90 ± 0.05	12.21 ± 0.05	13.79 ± 0.11
56 s / pH 8.0	9.46 ± 0.13	11.63 ± 0.10	13.22 ± 0.15
56 s / pH 5.0	9.46 ± 0.00	11.71 ± 0.00	13.32 ± 0.03

Figure 1

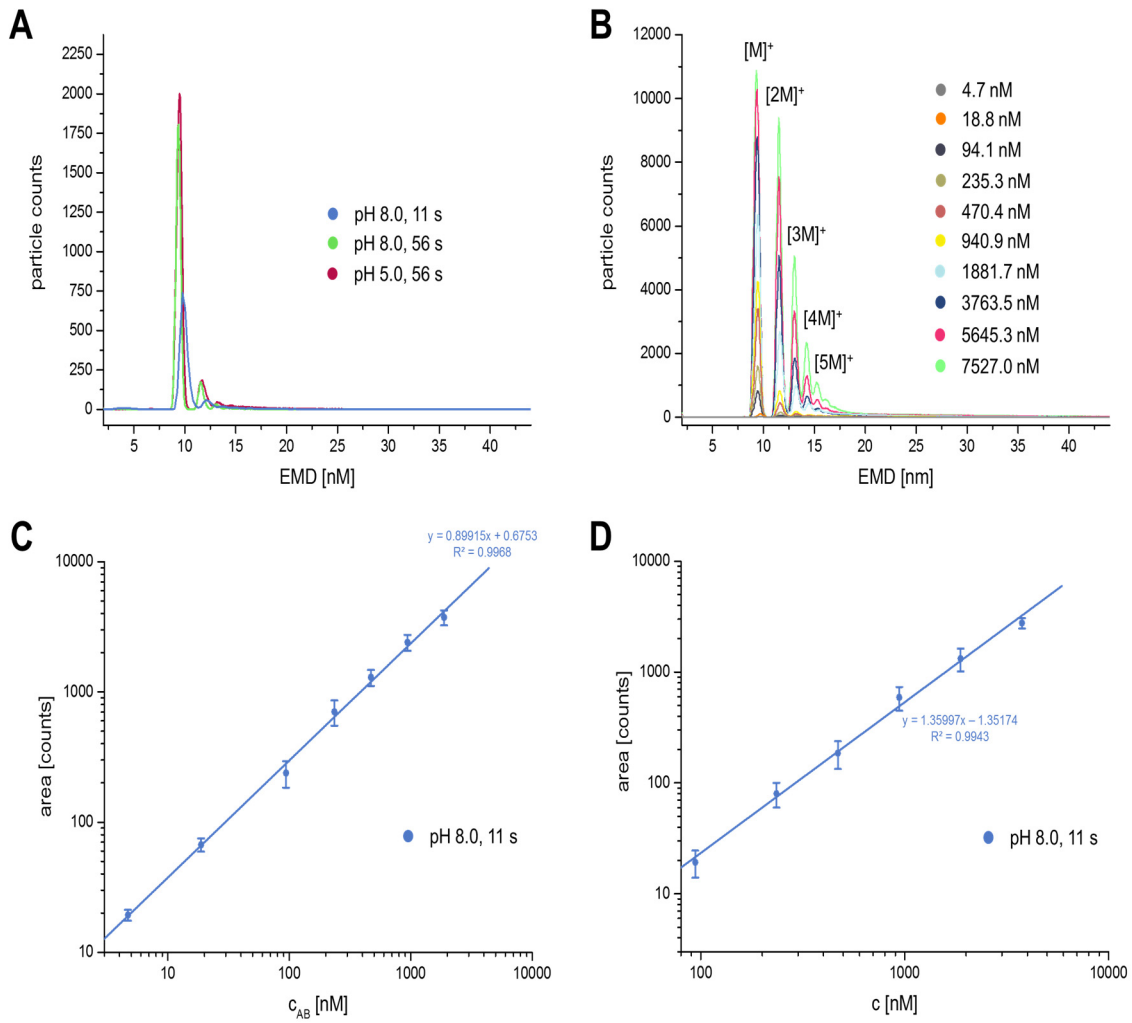


Figure 1: MacroIMS spectra of the recombinant monoclonal antibody (rmAB). 235.3 nM rmAB measured at three different conditions: pH 8.0 with 11 and 56 s DMA scan time and pH 5.0 with 56 s DMA scan time (**A**). An exemplary dilution series of rmAB ranging from 4.7 to 7527.0 nM measured at pH 8.0 and 56 s scan time (**B**). The obtained linear range based on peak areas (plot of the rmAB concentrations vs peak areas of the monomeric form of rmAB (**C**) and of the dimeric form of rmAB (**D**).

Figure 2

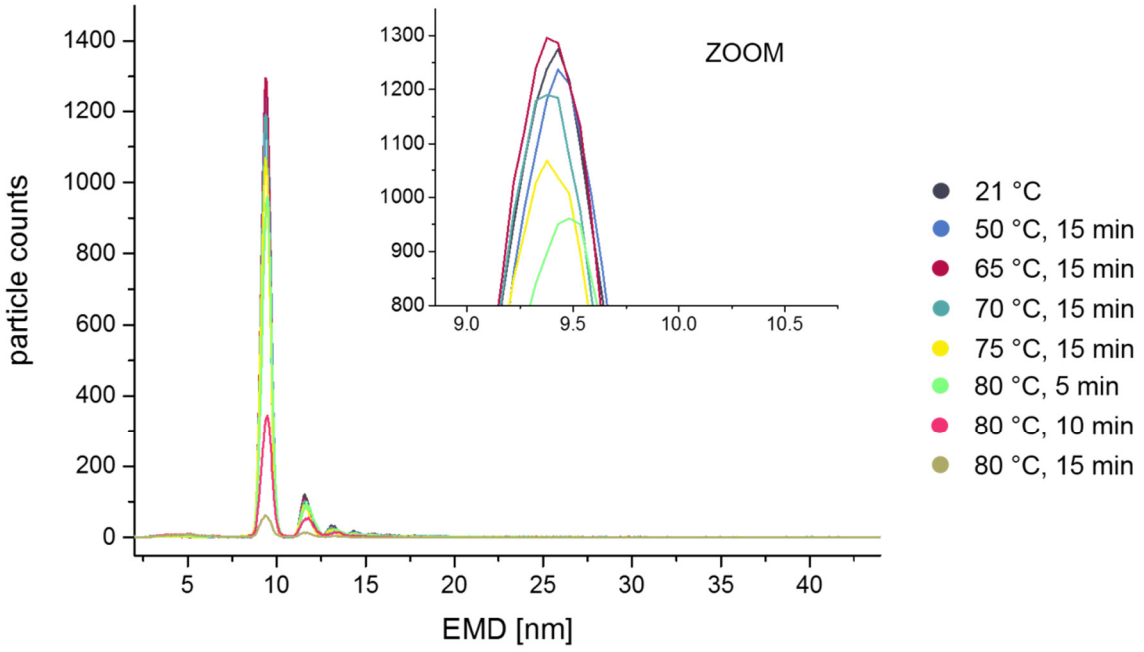


Figure 2: Influence of different temperatures and reaction times on 235.3 nM rmAB solutions directly monitored by MacroIMS with the Advanced Electrospray (Model 3482) device based on the soft X-ray charge conditioner. The inset shows the monomeric peak from RT (21 °C) up to 80 °C at 5 min.

Conclusions

The analysis of glycoproteins exhibits many challenges due to the huge structural and functional heterogeneity of the glycosylations and the fact that no universal analytical method can cover all aspects. Many techniques have to be combined for a complete understanding of glycoprotein structures, their occurrences, interactions, and functional purposes. This also includes appropriate strategies of isolating and purifying glycoproteins of interest. However, many methods lack some of the following points: sensitivity, robustness, or feasibility concerning ease of handling and time consumption.

The aim of the present thesis was to establish new rapid and reliable analytical approaches to qualitatively and quantitatively investigate intact glycoproteins in regard to size determination and interactions with lectins. Special focus was laid on electrophoretic methods in liquid- and gas-phase using MCGE and nES GEMMA, respectively.

For commercially available kits MCGE had previously been validated thoroughly only for unmodified proteins. As a consequence, the analysis of glycoproteins with this method required intensive evaluation of characteristic analytical parameters like accuracy and repeatability of sizing and quantification, as well as LODs and LOQs. As addressed in **Publication I**, two available protein assays (P230 and HSP-250) proved a good applicability for glycoproteins with high reproducibilities in sizing, high reliability and precision in quantitation, and high sensitivities (here the HSP-250 assay exceeded the P230 one). Only the accuracy of sizing showed restrictions with the determined molecular weights deviating from MS derived values. Moreover, these deviations clearly showed a correlation with the amount of glycosylation, i.e. with rising degrees of glycosylation, the deviations increasingly exceeded the MS values. These glycoprotein specific deviations were in accordance to their behaviour during SDS-PAGE analysis, yet to a higher extent with MCGE. This has also to do with the lack of well-defined glycosylated protein standards. However, due to the higher reliability in quantitation, its higher ease-of-use to the point of standardization, and particularly the savings in time and consumables, MCGE is a very good alternative to traditional SDS-PAGE. This might render it especially useful in routine analyses for process monitoring or quality control during production of biopharmaceuticals and biosimilars.

Publication II focussed on the targeted analysis of glycoproteins in complex biological samples. In this regard, MCGE was successfully combined with an affinity enrichment using lectin-coated magnetic beads. The combination of three lectins (Concanavalin A, wheat germ agglutinin, and *Sambucus nigra* agglutinin) with different and broader specificities allowed the enrichment of a wide

range of N-glycoproteins. The optimization of the enrichment step and the elution of the captured glycoproteins with the competitive sugars resulted in low unspecific binding of non-glycosylated proteins as shown by MCGE. Validation of the method with complex biological samples like human serum and *Trichoderma atroviride* mycelia extract proved the selectivity of the developed strategy. The results were confirmed by SDS-PAGE analysis. Additionally, two putative glycoproteins could be identified for the only fairly studied glycoproteome of the cellulose degrading fungus *Trichoderma atroviride*. However, also the existence of unspecific interactions between the biological samples and the magnetic beads themselves was revealed. Affecting glycoproteins of the sample, these interactions could increase the spectrum of enriched glycoproteins independently of the used lectin and thus influence specificity and lectin-glycoprotein interaction studies. If necessary, a detailed investigation of the unspecific binding by e.g. 2-D gel electrophoresis can be performed for further insights into the extent of impact on the specificity. At this point, this fact has to be considered for any data evaluation concerning a selective glycoprotein enrichment. Nevertheless, the developed strategy is a very sensitive and reliable approach for a selective glycoprotein analysis in complex biological samples. Although SDS-PAGE offers the possibility of subsequent mass spectrometric identification of the enriched analytes, MCGE benefits from time savings and lower sample consumption. Therefore, affinity enrichment combined with MCGE presents itself as a good alternative especially as a pre-analysis method before performing 2-D gel electrophoresis in quality control or research of unknown glycoproteomes.

Size-determinations of glycoproteins and the lectins were also performed in gas-phase with nES GEMMA as demonstrated in **Publication III**. MW derived from detected EMDs were in very good agreement with MALDI-MS values, and also the detected non-covalent glycoprotein-lectin complex values corroborated well with theoretically calculated approximations of an expected globular complex. This confirms the capability of the nES GEMMA system to keep even weak non-covalent interactions, which were formed in liquid-phase, intact throughout the nES process, separation, and detection. Moreover, it allowed direct interaction analyses of the glycoproteins and lectins employed during this work without prior separation of unbound glycoproteins and lectins, respectively. The experiments could be evaluated in regard to lectin binding specificities and strengths towards the glycoproteins. Results were compared to interaction studies analysed with CE-on-a-chip, which showed differences in binding affinities probably due to altered binding caused by fluorescent labelling of the glycoproteins.

Furthermore, the non-covalent biospecific complexes could be sampled onto nitrocellulose membrane and identified by a subsequent immunological assay. Additional sampling experiments with fluorescently labelled glycoproteins for subsequent fluorescent microscopy led to no positive results so far due to probably not enough analyte collection on the sampling surface. To overcome this

problem, sampling times longer than 36 hours on three consecutive days could eventually lead to a detectable signal based on rough calculations. However, longer sampling times affected the sample and the stability of the measurement. AFM measurements of the sampled glycoprotein-lectin complex were also inconclusive, as the size-differences between the interaction partners and the formed complex were too low for definite evaluations. In contrast, immunological results were irrevocable and proved, that the collected complex actually retained its biologically native structure throughout nES GEMMA analysis and sampling.

Consequently, nES GEMMA is a promising and straightforward platform especially for the analysis of weak interactions like that of glycoproteins and lectins, i.e. biospecific non-covalent complexes, which are difficult to maintain in vacuum-based instruments like mass spectrometers or TEM. It does not require a preceding separation of unreacted binding partners or elaborate sample preparation and scores with simplified data interpretation. Due to charge reduction to preferentially singly charged species, spectra are far less complex, i.e. straightforward in the interpretation, compared to ESI mass spectra of multiply charged non-covalent complexes (in native mass spectrometry) and data deconvolution processes can be omitted. Moreover, interactions can be directly assessed as they are formed in liquid-phase and, providing a label-free detection, binding is not altered or affected by covalent tags (e.g. fluorescent dye, biotin, or alkaline phosphatase) or coupling to solid supports as it is the case with for instance biosensors, ELLAs, or microarrays.

References

1. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* (**1999**), *1473*, 4-8.
2. Van den Steen P, Rudd PM, Dwek RA, Opdenakker G. Concepts and principles of O-linked glycosylation. *Crit Rev Biochem Mol Biol* (**1998**), *33*, 151-208.
3. Kornfeld R, Kornfeld S. Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* (**1985**), *54*, 631-64.
4. Burda P, Aebi M. The ALG10 locus of *Saccharomyces cerevisiae* encodes the alpha-1,2 glucosyltransferase of the endoplasmic reticulum: the terminal glucose of the lipid-linked oligosaccharide is required for efficient N-linked glycosylation. *Glycobiology* (**1998**), *8*, 455-62.
5. Spiro RG. Glucose residues as key determinants in the biosynthesis and quality control of glycoproteins with N-linked oligosaccharides. *J Biol Chem* (**2000**), *275*, 35657-60.
6. Johansen PG, Marshall RD, Neuberger A. Carbohydrates in protein. 3 The preparation and some of the properties of a glycopeptide from hen's-egg albumin. *Biochem J* (**1961**), *78*, 518-27.
7. Abbadi A, McHarfi M, Aubry A, Premilat S, Boussard G, Marraud M. Involvement of side functions in peptide structures: the Asx turn. Occurrence and conformational aspects. *J Am Chem Soc* (**1991**), *113*, 2729-35.
8. Imperiali B, Shannon KL, Unno M, Rickert KW. Mechanistic proposal for asparagine-linked glycosylation. *J Am Chem Soc* (**1992**), *114*, 7944-45.
9. Imperiali B. Protein glycosylation: the clash of the titans. *Acc Chem Res* (**1997**), *30*, 452-59.
10. Bause E, Legler G. The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr(Ser) for the N-glycosylation step during glycoprotein biosynthesis. *Biochem J* (**1981**), *195*, 639-44.
11. Bause E, Breuer W, Peters S. Investigation of the active site of oligosaccharyltransferase from pig liver using synthetic tripeptides as tools. *Biochem J* (**1995**), *312 (Pt 3)*, 979-85.
12. Helenius A, Aebi M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* (**2004**), *73*, 1019-49.
13. Dell A, Galadari A, Sastre F, Hitchen P. Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. *Int J Microbiol* (**2010**), *2010*, 148178.
14. Kowarik M, Young NM, Numao S, Schulz BL, Hug I, Callewaert N et al. Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J* (**2006**), *25*, 1957-66.
15. Rempe KA, Spruce LA, Porsch EA, Seeholzer SH, Norskov-Lauritsen N, St Geme JW, 3rd. Unconventional N-linked glycosylation promotes trimeric autotransporter function in *Kingella kingae* and *Aggregatibacter aphrophilus*. *MBio* (**2015**), *6*.
16. Fraser-Reid BO, Tatsuta K, Thiem J (eds). Glycoscience - chemistry and chemical biology. 2nd ed. Berlin: Springer; 2008.
17. Ito S, Yamashita K, Spiro RG, Kobata A. Structure of a carbohydrate moiety of a unit A glycopeptide of calf thyroglobulin. *J Biochem* (**1977**), *81*, 1621-31.
18. Herscovics A, Orlean P. Glycoprotein biosynthesis in yeast. *FASEB J* (**1993**), *7*, 540-50.
19. Stansell E, Desrosiers RC. Fundamental difference in the content of high-mannose carbohydrate in the HIV-1 and HIV-2 lineages. *J Virol* (**2010**), *84*, 8998-9009.
20. Wopereis S, Lefeber DJ, Morava E, Wevers RA. Mechanisms in protein O-glycan biosynthesis and clinical and molecular aspects of protein O-glycan biosynthesis defects: a review. *Clin Chem* (**2006**), *52*, 574-600.
21. Hounsell EF, Davies MJ, Renouf DV. O-linked protein glycosylation structure and function. *Glycoconj J* (**1996**), *13*, 19-26.
22. Hase S, Nishimura H, Kawabata S, Iwanaga S, Ikenaka T. The structure of (xylose)2glucose-O-serine 53 found in the first epidermal growth factor-like domain of bovine blood clotting factor IX. *J Biol Chem* (**1990**), *265*, 1858-61.
23. Harris RJ, Spellman MW. O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology* (**1993**), *3*, 219-24.

24. Shao L, Haltiwanger RS. O-fucose modifications of epidermal growth factor-like repeats and thrombospondin type 1 repeats: unusual modifications in unusual places. *Cell Mol Life Sci* (**2003**), *60*, 241-50.
25. Zachara NE, Hart GW. O-GlcNAc a sensor of cellular state: the role of nucleocytoplasmic glycosylation in modulating cellular function in response to nutrition and stress. *Biochim Biophys Acta* (**2004**), *1673*, 13-28.
26. Endo T. O-mannosyl glycans in mammals. *Biochim Biophys Acta* (**1999**), *1473*, 237-46.
27. Hart GW. Dynamic O-linked glycosylation of nuclear and cytoskeletal proteins. *Annu Rev Biochem* (**1997**), *66*, 315-35.
28. Gerwig GJ, Kamerling JP, Vliegenthart JF, Morag E, Lamed R, Bayer EA. Novel oligosaccharide constituents of the cellulase complex of *Bacteroides cellulosolvens*. *Eur J Biochem* (**1992**), *205*, 799-808.
29. Silbert JE, Bernfield M, Kokenyesi R. Proteoglycans: a special class of glycoproteins. In: Montreuil J, Schachter H, Vliegenthart JFG (eds.). New Comprehensive Biochemistry. Vol. 29b. Amsterdam: Elsevier; 1991. p. 1-31.
30. Häusler A, Ballou L, Ballou CE, Robbins PW. Yeast glycoprotein biosynthesis: MNT1 encodes an alpha-1,2-mannosyltransferase involved in O-glycosylation. *Proc Natl Acad Sci USA* (**1992**), *89*, 6846-50.
31. Faridmoayer A, Fentabil MA, Mills DC, Klassen JS, Feldman MF. Functional characterization of bacterial oligosaccharyltransferases involved in O-linked protein glycosylation. *J Bacteriol* (**2007**), *189*, 8088-98.
32. Musumeci MA, Faridmoayer A, Watanabe Y, Feldman MF. Evaluating the role of conserved amino acids in bacterial O-oligosaccharyltransferases by in vivo, in vitro and limited proteolysis assays. *Glycobiology* (**2014**), *24*, 39-50.
33. Varki A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* (**1993**), *3*, 97-130.
34. Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J (eds). Essentials of glycobiology. New York: Cold Spring Harbor Laboratory Press; 2009.
35. Coyne KE, Hall SE, Thompson S, Arce MA, Kinoshita T, Fujita T et al. Mapping of epitopes, glycosylation sites, and complement regulatory domains in human decay accelerating factor. *J Immunol* (**1992**), *149*, 2906-13.
36. O'Connor SE, Imperiali B. Conformational switching by asparagine-linked glycosylation. *J Am Chem Soc* (**1997**), *119*, 2295-6.
37. Meyer B, Möller H. Conformation of glycopeptides and glycoproteins. *Top Curr Chem* (**2007**), *267*, 187-251.
38. Shental-Bechor D, Levy Y. Folding of glycoproteins: toward understanding the biophysics of the glycosylation code. *Curr Opin Struct Biol* (**2009**), *19*, 524-33.
39. Banks DD. The effect of glycosylation on the folding kinetics of erythropoietin. *J Mol Biol* (**2011**), *412*, 536-50.
40. Tachibana Y, Fletcher GL, Fujitani N, Tsuda S, Monde K, Nishimura S. Antifreeze glycoproteins: elucidation of the structural motifs that are essential for antifreeze activity. *Angew Chem Int Ed Engl* (**2004**), *43*, 856-62.
41. Rudd PM, Joao HC, Coghill E, Fiten P, Saunders MR, Opdenakker G et al. Glycoforms modify the dynamic stability and functional activity of an enzyme. *Biochemistry* (**1994**), *33*, 17-22.
42. Elbein AD. The role of N-linked oligosaccharides in glycoprotein function. *Trends Biotechnol* (**1991**), *9*, 346-52.
43. Sharon N, Lis H. History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* (**2004**), *14*, 53R-62R.
44. Defaus S, Gupta P, Andreu D, Gutierrez-Gallego R. Mammalian protein glycosylation-structure versus function. *Analyst* (**2014**), *139*, 2944-67.
45. Lairson LL, Henrissat B, Davies GJ, Withers SG. Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* (**2008**), *77*, 521-55.
46. Watkins WM. The ABO blood group system: historical background. *Transfus Med* (**2001**), *11*, 243-65.

47. Landsteiner K. Über Agglunationserscheinungen normalen menschlichen Blutes. *Wein Klin Wschr* (**1901**), *14*, 1132-34.
48. Morgan WTJ, Watkins WM. The detection of a product of the blood group O gene and the relationship of the so-called O-substance to the agglutinogens A and B. *Br J Exp Pathol* (**1948**), *29*, 159-73.
49. Watkins WM, Morgan WT. Neutralization of the anti-H agglutinin in eel serum by simple sugars. *Nature* (**1952**), *169*, 825-6.
50. Morgan WTJ. The croonian lecture: a contribution to human biochemical genetics; the chemical basis of blood-group specificity. *Proc Roy Soc B* (**1960**), *151*, 308-47.
51. Ezzelarab M, Ayares D, Cooper DK. Carbohydrates in xenotransplantation. *Immunol Cell Biol* (**2005**), *83*, 396-404.
52. Kobayashi T, Cooper DK. Anti-Gal, alpha-Gal epitopes, and xenotransplantation. *Subcell Biochem* (**1999**), *32*, 229-57.
53. Morozumi K, Kobayashi T, Usami T, Oikawa T, Ohtsuka Y, Kato M et al. Significance of histochemical expression of Hanganutziu-Deicher antigens in pig, baboon and human tissues. *Transplant Proc* (**1999**), *31*, 942-4.
54. Sumner JB, Howell SF. Identification of hemagglutinin of jack bean with concanavalin A. *J Bacteriol* (**1936**), *32*, 227-37.
55. Boyd WC, Shapleigh E. Specific precipitating activity of plant agglutinins (lectins). *Science* (**1954**), *119*, 419.
56. Makela O. Studies in hemagglutinins of leguminosae seeds. *Ann Med Exp Biol Fenn* (**1957**), *35*, 1-133.
57. Sharon N, Lis H. Lectins: cell-agglutinating and sugar-specific proteins. *Science* (**1972**), *177*, 949-59.
58. Peumans WJ, Van Damme EJ, Barre A, Rouge P. Classification of plant lectins in families of structurally and evolutionary related proteins. *Adv Exp Med Biol* (**2001**), *491*, 27-54.
59. Kilpatrick DC. Animal lectins: a historical introduction and overview. *Biochim Biophys Acta* (**2002**), *1572*, 187-97.
60. Loris R. Principles of structures of animal and plant lectins. *Biochim Biophys Acta* (**2002**), *1572*, 198-208.
61. http://www.imperial.ac.uk/research/animallectins/ctld/lectins.html#Evolution Accessed 21.07.2016, 11:26 Uhr.
62. Yang RY, Rabinovich GA, Liu FT. Galectins: structure, function and therapeutic potential. *Expert Rev Mol Med* (**2008**), *10*, e17.
63. Drickamer K. C-type lectin-like domains. *Curr Opin Struct Biol* (**1999**), *9*, 585-90.
64. Montfort W, Villafranca JE, Monzingo AF, Ernst SR, Katzin B, Rutenber E et al. The three-dimensional structure of ricin at 2.8 A. *J Biol Chem* (**1987**), *262*, 5398-403.
65. DiGabriele AD, Lax I, Chen DI, Svahn CM, Jaye M, Schlessinger J et al. Structure of a heparin-linked biologically active dimer of fibroblast growth factor. *Nature* (**1998**), *393*, 812-7.
66. Wright CS. The crystal structure of wheat germ agglutinin at 2-2 A resolution. *J Mol Biol* (**1977**), *111*, 439-57.
67. Drenth J, Low BW, Richardson JS, Wright CS. The toxin-agglutinin fold. A new group of small protein structures organized around a four-disulfide core. *J Biol Chem* (**1980**), *255*, 2652-5.
68. Liang PH, Wang SK, Wong CH. Quantitative analysis of carbohydrate-protein interactions using glycan microarrays: determination of surface and solution dissociation constants. *J Am Chem Soc* (**2007**), *129*, 11177-84.
69. Syed P, Gidwani K, Kekki H, Leivo J, Pettersson K, Lamminmaki U. Role of lectin microarrays in cancer diagnosis. *Proteomics* (**2016**), *16*, 1257-65.
70. Weis WI, Drickamer K. Structural basis of lectin-carbohydrate recognition. *Annu Rev Biochem* (**1996**), *65*, 441-73.
71. Sacchettini JC, Baum LG, Brewer CF. Multivalent protein-carbohydrate interactions. A new paradigm for supermolecular assembly and signal transduction. *Biochemistry* (**2001**), *40*, 3009-15.

72. Schwefel D, Maierhofer C, Beck JG, Seeberger S, Diederichs K, Moller HM et al. Structural basis of multivalent binding to wheat germ agglutinin. *J Am Chem Soc* (**2010**), *132*, 8704-19.
73. Fuster MM, Esko JD. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat Rev Cancer* (**2005**), *5*, 526-42.
74. Staub A, Guillarme D, Schappler J, Veuthey JL, Rudaz S. Intact protein analysis in the biopharmaceutical field. *J Pharm Biomed Anal* (**2011**), *55*, 810-22.
75. Marino K, Bones J, Kattla JJ, Rudd PM. A systematic approach to protein glycosylation analysis: a path through the maze. *Nat Chem Biol* (**2010**), *6*, 713-23.
76. Pabst M, Altmann F. Glycan analysis by modern instrumental methods. *Proteomics* (**2011**), *11*, 631-43.
77. Yamada K, Kakehi K. Recent advances in the analysis of carbohydrates for biomedical use. *J Pharm Biomed Anal* (**2011**), *55*, 702-27.
78. Tharmalingam T, Adamczyk B, Doherty MA, Royle L, Rudd PM. Strategies for the profiling, characterisation and detailed structural analysis of N-linked oligosaccharides. *Glycoconj J* (**2013**), *30*, 137-46.
79. Reichel C, Thevis M. Gel electrophoretic methods for the analysis of biosimilar pharmaceuticals using the example of recombinant erythropoietin. *Bioanalysis* (**2013**), *5*, 587-602.
80. Rüchel R, Steere RL, Erbe EF. Transmission-electron microscopic observations of freeze etched polyacrylmaid gels. *J Chromatogr A* (**1978**), *166*, 563-75.
81. Mechref Y, Madera M, Novotny MV. Glycoprotein enrichment through lectin affinity techniques. *Methods Mol Biol* (**2008**), *424*, 373-96.
82. Arakawa T, Philo JS, Ejima D, Tsumoto K, Arisaka F. Aggregation analysis of therapeutic proteins, part 1. *BioProcess Int* (**2006**), *4*, 42-3.
83. Arakawa T, Philo JS, Ejima D, Tsumoto K, Arisaka F. Aggregation analysis of therapeutic proteins, part 2. *BioProcess Int* (**2007**), *5*, 36-47.
84. Arakawa T, Philo JS, Ejima D, Sato H, Tsumoto K. Aggregation analysis of therapeutic proteins, part 3 *BioProcess Int* (**2007**), *5*, 52-70.
85. Miller I, Crawford J, Gianazza E. Protein stains for proteomic applications: which, when, why? *Proteomics* (**2006**), *6*, 5385-408.
86. Furlan M, Perret BA, Beck EA. Staining of glycoproteins in polyacrylamide and agarose gels with fluorescent lectins. *Anal Biochem* (**1979**), *96*, 208-14.
87. Moroi M, Jung SM. Selective staining of human platelet glycoproteins using nitrocellulose transfer of electrophoresed proteins and peroxidase-conjugated lectins. *Biochim Biophys Acta* (**1984**), *798*, 295-301.
88. Thompson NJ, Rosati S, Rose RJ, Heck AJR. The impact of mass spectrometry on the study of intact antibodies: from post-translational modifications to structural analysis. *Chem Commun* (**2013**), *49*, 538-48.
89. Liu H, Zhang N, Wan D, Cui M, Liu Z, Liu S. Mass spectrometry-based analysis of glycoproteins and its clinical applications in cancer biomarker discovery. *Clin Proteomics* (**2014**), *11*, 14.
90. Banazadeh A, Veillon L, Wooding KM, Zabet-Moghaddam M, Mechref Y. Recent advances in mass spectrometric analysis of glycoproteins. *Electrophoresis* (**2017**), *38*, 162-89.
91. Wada Y, Dell A, Haslam SM, Tissot B, Canis K, Azadi P et al. Comparison of methods for profiling O-Glycosylation human proteome organisation human disease glycomics/proteome initiative multi-institutional study OF IgA1. *Mol Cell Proteomics* (**2010**), *9*, 719-27.
92. Yang Y, Liu F, Franc V, Halim LA, Schellekens H, Heck AJR. Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity. *Nat Commun* (**2016**), *7*, 13397.
93. Terral G, Beck A, Cianférani S. Insights from native mass spectrometry and ion mobility-mass spectrometry for antibody and antibody-based product characterization. *J Chromatogr B Analyt Technol Biomed Life Sci* (**2016**), *1032*, 79-90.
94. Hanisch FG. Top-down sequencing of O-glycoproteins by in-source decay matrix-assisted laser desorption ionization mass spectrometry for glycosylation site analysis. *Anal Chem* (**2011**), *83*, 4829-37.

95. Pedrali A, Tengattini S, Marrubini G, Bavaro T, Hemstrom P, Massolini G et al. Characterization of intact neo-glycoproteins by hydrophilic interaction liquid chromatography. *Molecules* (**2014**), *19*, 9070-88.
96. Rustandi RR. Hydrophobic interaction chromatography to analyze glycoproteins. *Methods Mol Biol* (**2013**), *988*, 211-9.
97. Mechref Y, Novotny MV. Structural investigations of glycoconjugates at high sensitivity. *Chem Rev* (**2002**), *102*, 321-69.
98. Landers JP (eds). Handbook of capillary and microchip electrophoresis and associated microtechniques. 3rd ed. Boca Raton: CRC Press; 2007.
99. Balaguer E, Demelbauer U, Pelzing M, Sanz-Nebot V, Barbosa J, Neususs C. Glycoform characterization of erythropoietin combining glycan and intact protein analysis by capillary electrophoresis - electrospray - time-of-flight mass spectrometry. *Electrophoresis* (**2006**), *27*, 2638-50.
100. Kato K, Yamaguchi Y. Glycoproteins and antibodies: solution NMR studies. eMagRes. Vol.: John Wiley & Sons, Ltd; 2007.
101. Dan X, Liu W, Ng TB. Development and applications of lectins as biological tools in biomedical research. *Med Res Rev* (**2016**), *36*, 221-47.
102. Loo D, Jones A, Hill MM. Lectin magnetic bead array for biomarker discovery. *J Proteome Res* (**2010**), *9*, 5496-500.
103. Choi E, Loo D, Dennis JW, O'Leary CA, Hill MM. High-throughput lectin magnetic bead array-coupled tandem mass spectrometry for glycoprotein biomarker discovery. *Electrophoresis* (**2011**), *32*, 3564-75.
104. Endo T. Fractionation of glycoprotein-derived oligosaccharides by affinity chromatography using immobilized lectin columns. *J Chromatogr A* (**1996**), *720*, 251-61.
105. Yang Z, Hancock WS. Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. *J Chromatogr A* (**2004**), *1053*, 79-88.
106. Monzo A, Bonn GK, Guttman A. Boronic acid-lectin affinity chromatography. 1. Simultaneous glycoprotein binding with selective or combined elution. *Anal Bioanal Chem* (**2007**), *389*, 2097-102.
107. Kasai K, Oda Y, Nishikata M, Ishii S. Frontal affinity chromatography: theory for its application to studies on specific interactions of biomolecules. *J Chromatogr* (**1986**), *376*, 33-47.
108. Ng ES, Yang F, Kameyama A, Palcic MM, Hindsgaul O, Schriemer DC. High-throughput screening for enzyme inhibitors using frontal affinity chromatography with liquid chromatography and mass spectrometry. *Anal Chem* (**2005**), *77*, 6125-33.
109. Tateno H, Nakamura-Tsuruta S, Hirabayashi J. Frontal affinity chromatography: sugar-protein interactions. *Nat Protoc* (**2007**), *2*, 2529-37.
110. Sato C, Yamakawa N, Kitajima K. Measurement of glycan-based interactions by frontal affinity chromatography and surface plasmon resonance. *Methods Enzymol* (**2010**), *478*, 219-32.
111. Ladbury JE. Application of isothermal titration calorimetry in the biological sciences: things are heating up! *Biotechniques* (**2004**), *37*, 885-7.
112. Christie MP, Toth I, Simerska P. Biophysical characterization of lectin-glycan interactions for therapeutics, vaccines and targeted drug-delivery. *Future Med Chem* (**2014**), *6*, 2113-29.
113. Ghai R, Falconer RJ, Collins BM. Applications of isothermal titration calorimetry in pure and applied research - survey of the literature from 2010. *J Mol Recognit* (**2012**), *25*, 32-52.
114. Dam TK, Brewer CF. Probing lectin-mucin interactions by isothermal titration microcalorimetry. *Methods Mol Biol* (**2015**), *1207*, 75-90.
115. Belicky S, Katrlik J, Tkac J. Glycan and lectin biosensors. *Essays Biochem* (**2016**), *60*, 37-47.
116. Liedberg B, Lundström I, Stenberg E. Principles of biosensing with an extended coupling matrix and surface plasmon resonance. *Sensor Actuat B: Chem* (**1993**), *11*, 63-72.
117. Janshoff A, Galla HJ, Steinem C. Piezoelectric mass-sensing devices as biosensors - an alternative to optical biosensors? *Angew Chem Int Ed Engl* (**2000**), *39*, 4004-32.
118. Gronewold TM. Surface acoustic wave sensors in the bioanalytical field: recent trends and challenges. *Anal Chim Acta* (**2007**), *603*, 119-28.

119. Chang BY, Park SM. Electrochemical impedance spectroscopy. *Annu Rev Anal Chem* (**2010**), *3*, 207-29.
120. McCoy JP, Jr., Varani J, Goldstein IJ. Enzyme-linked lectin assay (ELLA). II. Detection of carbohydrate groups on the surface of unfixed cells. *Exp Cell Res* (**1984**), *151*, 96-103.
121. Haab BB, Yue T. High-throughput studies of protein glycoforms using antibody-lectin sandwich arrays. *Methods Mol Biol* (**2011**), *785*, 223-36.
122. Thompson R, Creavin A, O'Connell M, O'Connor B, Clarke P. Optimization of the enzyme-linked lectin assay for enhanced glycoprotein and glycoconjugate analysis. *Anal Biochem* (**2011**), *413*, 114-22.
123. Gupta G, Surolia A, Sampathkumar SG. Lectin microarrays for glycomic analysis. *OMICS* (**2010**), *14*, 419-36.
124. Hirabayashi J, Kuno A, Tateno H. Lectin-based structural glycomics: a practical approach to complex glycans. *Electrophoresis* (**2011**), *32*, 1118-28.
125. Jorgenson JW, Lukacs KD. Capillary zone electrophoresis. *Science* (**1983**), *222*, 266-72.
126. McLaren DG, Chen DD. A quantitative study of continuous flow-counterbalanced capillary electrophoresis for sample purification. *Electrophoresis* (**2003**), *24*, 2887-95.
127. Altria KD. Fundamentals of capillary electrophoresis theory. *Methods Mol Biol* (**1996**), *52*, 3-13.
128. Bardelmeijer HA, Waterval J, Lingeman H, van't Hof R, Bult A, Underberg WJ. Pre-, on- and post-column derivatization in capillary electrophoresis. *Electrophoresis* (**1997**), *18*, 2214-27.
129. Underberg WJ, Waterval JC. Derivatization trends in capillary electrophoresis: an update. *Electrophoresis* (**2002**), *23*, 3922-33.
130. Dovichi NJ, Hu S. Chemical cytometry. *Curr Opin Chem Biol* (**2003**), *7*, 603-8.
131. Stoyanov AV, Ahmadzadeh H, Krylov SN. Heterogeneity of protein labeling with a fluorogenic reagent, 3-(2-furoyl)quinoline-2-carboxaldehyde. *J Chromatogr B Analyt Technol Biomed Life Sci* (**2002**), *780*, 283-7.
132. Wu D, Regnier FE. Sodium dodecyl sulfate-capillary gel electrophoresis of proteins using non-cross-linked polyacrylamide. *J Chromatogr* (**1992**), *608*, 349-56.
133. Nuchtavorn N, Smejkal P, Breadmore MC, Guijt RM, Doble P, Bek F et al. Exploring chip-capillary electrophoresis-laser-induced fluorescence field-deployable platform flexibility: separations of fluorescent dyes by chip-based non-aqueous capillary electrophoresis. *J Chromatogr A* (**2013**), *1286*, 216-21.
134. Bousse L, Mouradian S, Minalla A, Yee H, Williams K, Dubrow R. Protein sizing on a microchip. *Anal Chem* (**2001**), *73*, 1207-12.
135. Kuschel M, Neumann T, Barthmaier P, Kratzmeier M. Use of lab-on-a-chip technology for protein sizing and quantitation. *J Biomol Tech* (**2002**), *13*, 172-8.
136. Chow AW. Protein separations. *Methods Mol Biol* (**2006**), *339*, 145-58.
137. Ramsey JM (inventor) Apparatus and method for performing microfluidic manipulations for chemical analysis and synthesis. US patent 6010607A (2000).
138. Bousse L, Dubrow B, Ulfelder K. High-performance DNA separations in microchipelectrophoresis systems. In: Harrison DJ, van den Berg A (eds.). Micro total analysis systems '98. Vol. 1. Dordrecht: Springer Science+Business Media; 1998.
139. Harvey MD, Bandilla D, Banks PR. Subnanomolar detection limit for sodium dodecyl sulfate-capillary gel electrophoresis using a fluorogenic, noncovalent dye. *Electrophoresis* (**1998**), *19*, 2169-74.
140. Wenz C, Marchetti-Deschmann M, Herwig E, Schrottner E, Allmaier G, Trojer L et al. A fluorescent derivatization method of proteins for the detection of low-level impurities by microchip capillary gel electrophoresis. *Electrophoresis* (**2010**), *31*, 611-7.
141. Grunert T, Marchetti-Deschmann M, Miller I, Müller M, Allmaier G. Comparing the applicability of CGE-on-the-chip and SDS-PAGE for fast pre-screening of mouse serum samples prior to proteomics analysis. *Electrophoresis* (**2008**), *29*, 4332-40.
142. Weiss VU, Kolivoska V, Kremser L, Gas B, Blaas D, Kenndler E. Virus analysis by electrophoresis on a microfluidic chip. *J Chromatogr B Analyt Technol Biomed Life Sci* (**2007**), *860*, 173-9.
143. Hsieh JF, Chen ST. Comparative studies on the analysis of glycoproteins and lipopolysaccharides by the gel-based microchip and SDS-PAGE. *Biomicrofluidics* (**2007**), *1*, 14102.

144. Kelly L, Barthmaier P. Glycoprotein sizing on the Agilent 2100 bioanalyzer. *Agilent Application Note 5989-0332EN* (**2007**).
145. Kaufman SL, Skogen JW, Dorman FD, Zarrin F, Lewis KC. Macromolecule analysis based on electrophoretic mobility in air: globular proteins. *Anal Chem* (**1996**), *68*, 1895-904.
146. Kaufman SL. Analysis of biomolecules using electrospray and nanoparticle methods: the gas-phase electrophoretic mobility molecular analyzer (GEMMA). *J Aerosol Sci* (**1998**), *29*, 537-52.
147. de la Mora JF, de Juan L, Eichler T, Rosell J. Differential mobility analysis of molecular ions and nanometer particles. *TrAC, Trends Anal Chem* (**1998**), *17*, 328-39.
148. Bacher G, Szymanski WW, Kaufman SL, Zollner P, Blaas D, Allmaier G. Charge-reduced nano electrospray ionization combined with differential mobility analysis of peptides, proteins, glycoproteins, noncovalent protein complexes and viruses. *J Mass Spectrom* (**2001**), *36*, 1038-52.
149. Mouradian S, Skogen JW, Dorman FD, Zarrin F, Kaufman SL, Smith LM. DNA analysis using an electrospray scanning mobility particle sizer. *Anal Chem* (**1997**), *69*, 919-25.
150. Kaddis CS, Loo JA. Native protein MS and ion mobility: large flying proteins with ESI. *Anal Chem* (**2007**), *79*, 1778-84.
151. Saucy DA, Ude S, Lenggoro IW, Fernandez de la Mora J. Mass analysis of water-soluble polymers by mobility measurement of charge-reduced ions generated by electrosprays. *Anal Chem* (**2004**), *76*, 1045-53.
152. Müller R, Laschober C, Szymanski WW, Allmaier G. Determination of molecular weight, particle size, and density of high number generation PAMAM dendrimers using MALDI−TOF−MS and nES−GEMMA. *Macromolecules* (**2007**), *40*, 5599-605.
153. Kemptner J, Marchetti-Deschmann M, Siekmann J, Turecek PL, Schwarz HP, Allmaier G. GEMMA and MALDI-TOF MS of reactive PEGs for pharmaceutical applications. *J Pharm Biomed Anal* (**2010**), *52*, 432-7.
154. Pease LF, 3rd. Physical analysis of virus particles using electrospray differential mobility analysis. *Trends Biotechnol* (**2012**), *30*, 216-24.
155. Laschober C, Wruss J, Blaas D, Szymanski WW, Allmaier G. Gas-phase electrophoretic molecular mobility analysis of size and stoichiometry of complexes of a common cold virus with antibody and soluble receptor molecules. *Anal Chem* (**2008**), *80*, 2261-4.
156. Pease LF, 3rd, Lipin DI, Tsai DH, Zachariah MR, Lua LH, Tarlov MJ et al. Quantitative characterization of virus-like particles by asymmetrical flow field flow fractionation, electrospray differential mobility analysis, and transmission electron microscopy. *Biotechnol Bioeng* (**2009**), *102*, 845-55.
157. Pease LF, Tsai D-H, Zangmeister RA, Zachariah MR, Tarlov MJ. Quantifying the surface coverage of conjugate molecules on functionalized nanoparticles. *J Phys Chem C* (**2007**), *111*, 17155-7.
158. Tsai DH, DelRio FW, Keene AM, Tyner KM, MacCuspie RI, Cho TJ et al. Adsorption and conformation of serum albumin protein on gold nanoparticles investigated using dimensional measurements and in situ spectroscopic methods. *Langmuir* (**2011**), *27*, 2464-77.
159. Loo JA, Berhane B, Kaddis CS, Wooding KM, Xie Y, Kaufman SL et al. Electrospray ionization mass spectrometry and ion mobility analysis of the 20S proteasome complex. *J Am Soc Mass Spectrom* (**2005**), *16*, 998-1008.
160. de la Mora JF, Ude S, Thomson BA. The potential of differential mobility analysis coupled to MS for the study of very large singly and multiply charged proteins and protein complexes in the gas phase. *Biotechnol J* (**2006**), *1*, 988-97.
161. Kaddis CS, Lomeli SH, Yin S, Berhane B, Apostol MI, Kickhoefer VA et al. Sizing large proteins and protein complexes by electrospray ionization mass spectrometry and ion mobility. *J Am Soc Mass Spectrom* (**2007**), *18*, 1206-16.
162. Pease LF, 3rd, Elliott JT, Tsai DH, Zachariah MR, Tarlov MJ. Determination of protein aggregation with differential mobility analysis: application to IgG antibody. *Biotechnol Bioeng* (**2008**), *101*, 1214-22.
163. Guha S, Wayment JR, Tarlov MJ, Zachariah MR. Electrospray-differential mobility analysis as an orthogonal tool to size-exclusion chromatography for characterization of protein aggregates. *J Pharm Sci* (**2012**), *101*, 1985-94.

164. Taylor G. Disintegration of water drops in electric field. *Proc R Soc Lond A Math Phys Sci* (**1964**), *280*, 383-97.
165. Cloupeau M, Prunet-Foch B. Electrostatic spraying of liquids in cone-jet mode. *J Electrostat* (**1989**), *22*, 135-59.
166. Model 3980C macroIMS™ macroion mobility spectrometer - operation and service manual *P/N 1930110, Revision B* (**2008**).
167. Fuchs NA. On the stationary charge distribution on aerosol particles in a bipolar ionic atmosphere. *Geofis Pura Appl* (**1963**), *56*, 185-93.
168. Wiedensohler A. An approximation of the bipolar charge distribution for particles in the sub-micron size range. *J Aerosol Sci* (**1988**), *19*, 387-9.
169. Laschober C, Kaufman SL, Reischl G, Allmaier G, Szymanski WW. Comparison between an unipolar corona charger and a polonium-based bipolar neutralizer for the analysis of nanosized particles and biopolymers. *J Nanosci Nanotechnol* (**2006**), *6*, 1474-81.
170. Kallinger P, Szymanski WW. Experimental determination of the steady-state charging probabilities and particle size conservation in non-radioactive and radioactive bipolar aerosol chargers in the size range of 5–40 nm. *J Nanopart Res* (**2015**), *17*, 171.
171. Iribarne JV, Thomson BA. On the evaporation of small ions from charged droplets. *J Chem Phys* (**1976**), *64*, 2287-94.
172. Thomson BA, Iribarne JV. Field induced ion evaporation from liquid surfaces at atmospheric pressure. *J Chem Phys* (**1979**), *71*, 4451-63.
173. Konermann L, Rodriguez AD, Liu J. On the formation of highly charged gaseous ions from unfolded proteins by electrospray ionization. *Analytical Chemistry* (**2012**), *84*, 6798-804.
174. Ahadi E, Konermann L. Modeling the behavior of coarse-grained polymer chains in charged water droplets: implications for the mechanism of electrospray ionization. *J Phys Chem B* (**2012**), *116*, 104-12.
175. Dole M, Mack LL, Hines RL, Mobley RC, Ferguson LD, Alice MB. Molecular beams of macroions. *J Chem Phys* (**1968**), *49*, 2240-9.
176. Konermann L, Ahadi E, Rodriguez AD, Vahidi S. Unraveling the mechanism of electrospray ionization. *Anal Chem* (**2013**), *85*, 2-9.
177. Kaufman SL. Electrospray diagnostics performed by using sucrose and proteins in the gas-phase electrophoretic mobility molecular analyzer (GEMMA). *Anal Chim Acta* (**2000**), *406*, 3-10.
178. Lewis KC, Dohmeier DM, Jorgenson JW, Kaufman SL, Zarrin F, Dorman FD. Electrospray-condensation particle counter: a molecule-counting LC detector for macromolecules. *Anal Chem* (**1994**), *66*, 2285-92.
179. Li M, Guha S, Zangmeister R, Tarlov MJ, Zachariah MR. Quantification and compensation of nonspecific analyte aggregation in electrospray sampling. *Aerosol Sci Technol* (**2011**), *45*, 849-60.
180. Flagan RC. History of electrical aerosol measurements. *Aerosol Sci Technol* (**1998**), *28*, 301-80.
181. Tammet H. Size and mobility of nanometer particles, clusters and ions. *Aerosol Sci* (**1995**), *26*, 459-75.
182. Knutson EO, Whitby KT. Aerosol classification by electric mobility: apparatus theory and applications. *Aerosol Sci* (**1975**), *6*, 443-51.
183. Lewis KC, Dohmeier DM, Jorgenson JW, Kaufman SL, Zarrin F, Dorman FD. Electrospray-condensation particle counter: a molecule-counting LC detector for macromolecules. *Anal Chem* (**1994**), *66*, 2285-92.
184. Sem GJ. Design and performance characteristics of three continuous-flow condensation particle counters: a summary. *Atmos Res* (**2002**), *62*, 267-94.
185. McDermott WT, Ockovic RC, Stolzenburg MR. Counting efficiency of an improved 30-Å condensation nucleus counter. *aerosol Sci Technol* (**1991**), *14*, 278-87.
186. Kallinger P, Weiss VU, Lehner A, Allmaier G, Szymanski WW. Analysis and handling of bio-nanoparticles and environmental nanoparticles using electrostatic aerosol mobility. *Particuology* (**2013**), *11*, 14-9.
187. Fissan H, Kennedy MK, Krinke TJ, Kruis FE. Nanoparticles from the gas phase as building blocks for electrical devices. *J Nanopart Res* (**2003**), *5*, 299-310.

188. Havlik M, Marchetti-Deschmann M, Friedbacher G, Winkler W, Messner P, Perez-Burgos L et al. Comprehensive size-determination of whole virus vaccine particles using gas-phase electrophoretic mobility macromolecular analyzer, atomic force microscopy, and transmission electron microscopy. *Anal Chem* (**2015**), *87*, 8657-64.

Abbreviations

AFM	atomic force microscopy
Asn	asparagine
Asp	aspartic acid
BGE	background electrolyte
CBP	carbohydrate-binding protein
CE	capillary electrophoresis
CEM	chain ejection model
CGE	capillary gel electrophoresis
CIEF	capillary isoelectric focussing
cmc	critical micelle concentration
ConA	concanavalin A
CPC	condensation particle counter
CRM	charge-residue model
CZE	capillary zone electrophoresis
DLS	dynamic light scattering
DTT	dithiotreitol
EIS	electrochemical impedance spectroscopy
ELLA	enzyme-linked lectin assay
ELISA	enzyme-linked immunosorbent assay
EMD	electrophoretic mobility analyzer
ENAS	electrostatic nanoparticle sampler
EOF	electroosmotic flow
EPO	erythropoietin
ER	endoplasmic reticulum
ESI	electrospray ionization
FAC	frontal affinity chromatography
FFF	field-flow fractionation
FT-ICR	Fourier-transform ion cyclotron resonance
Fuc	fucose
FucT	fucosyltransferase
GE	gel electrophoresis
Gal	galactose

GalT	galactosyltransferase
GalNAc	N-acetylgalactosamine
GalNAcT	N-acetylgalactosaminyltransferase
GEMMA	gas-phase electrophoretic mobility molecular analyzer
Glc	glucose
GlcNAc	N-acetylglucosamine
HILIC	hydrophilic interaction liquid chromatography
HSP-250	High Sensitivity Protein 250 Kit
Hyl	hydroxylysine
Hyp	hydroxyproline
IEF	isoelectric focussing
IEM	ion evaporation model
IEX	ion exchange chromatography
IMS	ion mobility spectrometry
ITC	isothermal titration calorimetry
LC	liquid chromatography
LIF	laser-induced fluorescence
LOD	limit of detection
LOQ	limit of quantification
MALDI	matrix-assisted laser desorption/ionization
Man	manose
MCGE	microchip capillary gel electrophoresis
MEKC	micellar electrokinetic chromatography
Met	methionine
MLAC	mixed-bed or multi-lectin affinity chromatography
MS	mass spectrometry
nDMA	nano differential mobility analyzer
nES	nano electrospray
NHS	N-hydroxy succinimide
OT	oligosaccharyltransferase
P230	Protein 230 Kit
PAGE	polyacrylamide gel electrophoresis
Pro	proline
Q	quadrupole
QCM	quartz crystal microbalance
RNA	ribonucleic acid

RNase	ribonuclease
RP	reversed-phase
SAM	self-assembled monolayer
SAW	surface acoustic wave
SDS	sodium dodecyl sulfate
SEC	size-exclusion chromatography
Ser	serine
SLAC	serial lectin affinity chromatography
SNA	*Sambucus nigra* agglutinin
SPR	surface plasmon resonance
TEM	transmission electron microscopy
Thr	threonine
TOF	time-of-flight
Tyr	tyrosine
Val	valine
WGA	wheat germ agglutinin
Xyl	xylose

Curriculum Vitae

Name:	**Nicole Yvonne Engel**
Date of birth:	**17.01.1985**
Place of birth:	**Erding (Germany)**
Nationality:	**German**

Education

01/2012 – present	**PhD studies, TU Wien (Austria)** • Institute for Chemical Technologies and Analytics, research group Bio and Polymer Analysis, Prof. Dr. Günter Allmaier • PhD thesis: „*Electrophoretic analysis of glycoproteins and their interactions with lectins studied by means of microchip capillary electrophoresis in the liquid phase and of gas-phase electrophoretic molecular mobility analyzer at atmospheric pressure*“
04/2009 – 03/2011	**Master of Science (M.Sc.), University of Konstanz (Germany)** • area of study: chemistry • Master thesis: „*Protein Interaction Analysis by High Resolution Bioaffinity Mass Spectrometry*“ (Prof. Dr. Dr. Michael Przybylski, research group Analytical Chemistry)
10/2005 – 10/2008	**Bachelor of Science (B.Sc.), University of Konstanz (Germany)** • area of study: chemistry • Bachelor thesis: „*Neue Taq DNA-Polymerase für die Real-Time RT-PCR-Analyse*“ (Prof. Dr. Andreas Marx, research group Organic Chemistry)
09/1995 – 06/2004	**Abitur, Gymnasium Erding (Germany)**

Work experience

01/2012 – 01/2017	**TU Wien (Austria)** • project assistant (Institute for Chemical Technologies and Analytics, research group Bio and Polymer Analysis, Prof. Dr. Günter Allmaier)
11/2008 – 02/2009	**Hexal AG, Department Bioanalytics (Oberhaching, Germany)** • internship • focus: establishment, optimization, and validation of ELISAs in consideration of GLP standards
09/2004 – 12/2004	**Airport Munich (Germany**) • airport security representative
07/2004 – 09/2004	**Haberl Gastronomie GmbH & Co. (Munich, Germany)** • temporary job

Publications

R. Kranaster, M. Drum, **N. Engel**, M. Weidmann, F. T. Hufert, A. Marx: *One-step RNA pathogen detection with reverse transcriptase activity of a mutated thermostable Thermus aquaticus DNA polymerase*, Biotechnol. J. (2010), *2*, 224-31.

N. Engel, V. U. Weiss, C. Wenz, A. Rüfer, M. Kratzmeier, S. Glück, M. Marchetti-Deschmann, G. Allmaier: *A comparative study of analytical parameters for proteins with different degrees of glycosylation*, Agilent Application Note (2013), 5991-3435EN.

S. Slamnoiu, C. Vlad, M. Stumbaum, A. Moise, K. Lindner, **N. Engel**, M. Vilanova, M. Diaz, C. Karreman, M. Leist, T. Ciossek, B. Hengerer, M. Vilaseca, M. Przybylski: *Identification and Affinity-Quantification of ß-Amyloid and α-Synuclein Polypeptides Using On-Line SAW-Biosensor-Mass Spectrometry*, JASMS (2014), *25*, 1472-81.

N. Engel, V. U. Weiss, C. Wenz, A. Rüfer, M. Kratzmeier, S. Glück, M. Marchetti-Deschmann, G. Allmaier: *Challenges of glycoprotein analysis by microchip capillary gel electrophoresis*, Electrophoresis (2015), *36*, 1754-58.

N. Y. Engel, V. U. Weiss, M. Marchetti-Deschmann, G. Allmaier: *nES GEMMA analysis of lectins and their interactions with glycoproteins – separation, detection, and sampling of non-covalent biospecific complexes*, JASMS (2017), *28*, 77-86.

N. Y. Engel, V. U. Weiss, N. Puffler, G. Allmaier: *MacroIMS Analysis of a Recombinant Monoclonal Antibody – Influences of Scan Time, Electrolyte pH, and Temperature Treatment of Sample on Obtained Signals*, TSI Application Note, submitted 2016.

N. Y. Engel, V. U. Weiss, C. Wenz, S. Glück, A. Rüfer, M. Kratzmeier, M. Marchetti-Deschmann, G. Allmaier: *Microchip capillary gel electrophoresis combined with lectin affinity enrichment employing magnetic beads for glycoprotein analysis*, Anal. Bioanal. Chem., to be submitted 2017.

Contributions to conferences

N. Engel, V. U. Weiss, M. Kratzmaier, S. Glück, A. Ruefer, M. Marchetti-Deschmann, G. Allmaier: "Glycoprotein affinity enrichment combined with analysis by microchip capillary gel electrophoresis (MCGE)"; 9. ASAC JunganalytikerInnenforum, Vienna, Austria, 06/2013. (talk)

D. Resetar, M. Klobucar, S. Kralevic Pavelic, **N. Engel**, M. Marchetti-Deschmann, M. Staver, J. Peter Katalinic: "High throughput analytical platforms for Adriatic-brand food quality and authenticity - Croatian dry-cured ham analysis"; 7th MSBM (Mass Spectrometry in Biotechnology and Medicine), Dubrovnik, Croatia, 07/2013. (poster)

N. Engel, C. Wenz, A. Ruefer, M. Kratzmaier, S. Glück, M. Marchetti-Deschmann, G. Allmaier: "Lectin affinity enrichment in combination with microchip capillary gel electrophoresis

(MCGE) for sensitive and selective glycoprotein analysis"; 30th International Symposium on MicroScale Bioseparations, Pecs, Hungary, 04/2014. (talk)

N. Engel, V. U. Weiss, M. Marchetti-Deschmann, G. Allmaier: "Challenges in lectin affinity enrichment for subsequent selective glycoprotein analysis in 1D- and 2D-GE"; 8th Central and Eastern European Proteomics Conference, Vienna, Austria, 07/2014. (poster and poster talk)

N. Engel, V. U. Weiss, M. Marchetti-Deschmann, G. Allmaier: "Pitfalls of a selective glycoprotein enrichment demonstrated by microchip-capillary-gel-electrophoresis"; ANAKON 2015, Graz, Austria, 03/2015. (poster)

N. Engel, V. U. Weiss, M. Marchetti-Deschmann, G. Allmaier: "Lectin-glycoprotein interaction analysis with MacroIMS“; 1st TSI Seminar at TU Wien, Vienna, Austria, 06/2015. (talk)

N. Engel, V. U. Weiss, M. Marchetti-Deschmann, G. Allmaier: "Separation, detection and collection of non-covalent biospecific lectin-glycoprotein complexes with a nanoelectrospray gas-phase electrophoretic mobility molecular analyzer (nES GEMMA)"; 31st Asimolar Conference on Mass Spectrometry - Native mass spectrometry-based structural biology, Asimolar Conference Center, Pacific Grove (CA), USA, 10/2015. (poster and poster highlight talk)

N. Y. Engel: "Challenges of glycoprotein affinity enrichment using lectins analyzed by means of MCGE, SDS-PAGE and GEMMA"; seminar "Moderne Analytische Chemie", TU Wien, Vienna, Austria (invited), 11/2015. (talk)

N. Y. Engel, V. U. Weiss, M. Marchetti-Deschmann, G. Allmaier: “Lectin-glycoprotein interaction analysis by means of nano electrospray gas-phase electrophoretic mobility molecular analyzer (GEMMA)”; 34th Informal Meeting on Mass Spectrometry, Fiera di Primiero, Italy, 05/2016. (talk)