Multi-level structural and functional characterization of therapeutic glycoproteins by mass spectrometry

Lippold, S.

MULTI-LEVEL STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THERAPEUTIC GLYCOPROTEINS BY MASS SPECTROMETRY

ISBN 978-94-6419-354-1

The work presented in this thesis was performed at the Center for Proteomics and Metabolomics, Leiden University Medical Center, The Netherlands.

The work was supported by the European Union Horizon 2020 Program Analytics for Biologics, grant number 765502.

Cover design: Lennart Brauer

Printing: Gildeprint

MULTI-LEVEL STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THERAPEUTIC GLYCOPROTEINS BY MASS SPECTROMETRY

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van de rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties,
te verdedigen op dinsdag 30 november 2021
klokke 15:00 uur

door

Steffen LIPPOLD

Geboren te Goslar, Duitsland
op 19 april 1992

Promotor:	Prof. dr. M. Wuhrer
Co-promotors:	Dr. N. de Haan Dr. D. Falck
Leden promotiecomissie:	Prof.dr. R.E.M. Toes Prof. dr. R.P.H. Bischoff, *Department of Analytical Biochemistry,* *University of Groningen, Groningen,* *The Netherlands* Prof. dr. F. Lisacek, *Proteome Informatics Group,* *Swiss Institute of Bioinformatics, Geneva,* *Switzerland* Dr. K.R. Reiding, *Bijvoet Center for Biomolecular Research* *and Utrecht Institute for Pharmaceutical Sciences,* *University of Utrecht, Utrecht* *The Netherlands*

CONTENTS

1

INTRODUCTION

1.1. THERAPEUTIC PROTEINS

THERAPEUTIC proteins are successful and important biopharmaceuticals and many of them are on the World Health Organization list of essential medicines [1, 2]. Recombinant protein-based medicines have been introduced in recent decades, improving accessibility and enabling new treatment opportunities. Important milestones were the approval of recombinant human insulin (Humulin®, 1982), recombinant human erythropoietin (Epogen®, 1989) and the first monoclonal antibody (mAb) for targeted cancer therapy (Rituxan®, 1997) by the Food and Drug Administration (FDA) [3–5]. Nowadays, mAbs are the dominating entity of therapeutic proteins and have been established for therapeutic areas such as oncology, dermatology, hematology and neurology [1]. In addition, more complex formats, e.g. bispecific antibodies or antibody-drug conjugates, are currently emerging [6]. Numerous mAb-based drugs are on the market (100 FDA approvals, status April 2021) with a steadily increasing number in development [6, 7]. Recombinant therapeutic proteins are much larger compared to small molecule (synthetic) drugs and exhibit an inherently higher degree of structural complexity.

1.1.1. GLYCOSYLATION OF THERAPEUTIC PROTEINS

Glycosylation of (therapeutic) proteins is of high importance for their physico-chemical properties, such as solubility and stability, as well as for pharmacokinetics and pharmacodynamics [8, 9]. For example, glycosylation affects the half-life of erythropoietin (EPO) [10, 11]. Glycosylation also impacts the safety and efficacy of therapeutic proteins [12]. A protein may vary in the occupancy of a glycosylation site (macroheterogeneity) or in the glycan structure present at a single glycosylation site (microheterogeneity) [13]. Protein glycosylation is categorized into *N*- and *O*-glycosylation (**Figure 1.1**). Asn residues in a consensus sequence motive (Asn-x-Ser/Thr, x ≠ Pro) may be subjected to *N*-glycosylation [14]. In contrast, *O*-glycosylation (Ser/Thr) cannot be easily predicted by a specific sequence motif [15].

The most common building blocks of mammalian protein glycans are hexoses (galactose, mannose), *N*-acetylhexosamines (*N*-acetylglucosamine, *N*-acetylgalactosamine), deoxyhexose (fucose) and sialic acids (*N*-acetylneuraminic acid and *N*-glycolylneuraminic acid) (**Figure 1.1**). Furthermore, glycan building blocks may be decorated by other modifications such as acetylation or phosphorylation. *N*-glycans share a common core and are structurally divided into high mannose-, hybrid- and complex-type glycans. Mucin-type *O*-glycans show an *N*-acetylgalactosamine starting unit and are grouped according to eight different core structures [16]. On therapeutic proteins, the most often found *O*-glycans are of the core 1 type.

The glycosylation features of recombinant therapeutic proteins are highly dependent on the production host [17]. For therapeutic proteins relying on glycosylation for the mechanism of action, mammalian cell lines are usually chosen. The glycosylation profiles of mammalian cell derived proteins are closer to human glycosylation

compared to bacterial or fungal production systems. Common cell lines for recombinant therapeutic protein production are Chinese hamster ovary (CHO), baby hamster kidney (BHK), murine myeloma cells (NS0, SP2/0) or human embryonic kidney 293 (HEK293) cells [18, 19]. Of note, hamster and murine cells additionally express non-human glycan features which negatively affect the safety of therapeutic proteins. For example, the presence of *N*-glycolylneuraminic acid and α1-3-linked digalactose motives may cause immunogenicity [20].

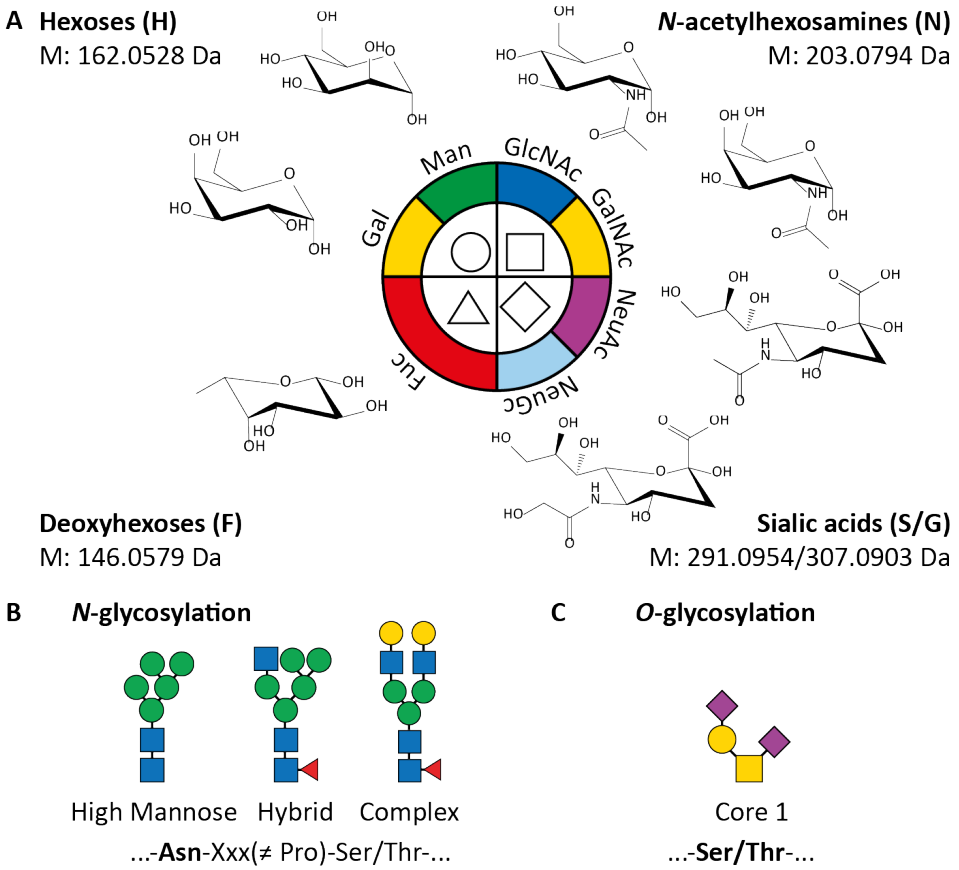

Figure 1.1: Common glycan building blocks of therapeutic protein glycosylation. **A** Monosaccharides are depicted with α-chair structure, symbol and color-coding. Monoisotopic masses (M) of monosaccharide residuals are indicated. **B** and **C** Representation of common *N*- and *O*-glycan structures observed for therapeutic proteins. Of note, all information is related to mammalian-based expression systems, which are usually used for therapeutic protein production.

1.1.2. STRUCTURE OF IMMUNOGLOBULIN G-BASED ANTIBODIES

The protein sequence of most therapeutic mAbs is based on human immunoglobulin G (IgG) [21]. IgG consist of two heavy and two light chains, arranged in a Y-form and connected via disulfide bonds (**Figure 1.2**). The antigen-binding fragment (Fab) is formed by the Fd (VH and CH1) and the light chain (VL and CL). Three complementarity-determining regions (CDR) in the Fab are responsible for antigen binding. The Fab is connected to the fragment crystallizable (Fc) domain (CH2 and CH3) via a hinge region. IgG subclasses 1, 2 and 4 are currently being used for therapeutic mAbs. The subclasses differ in specific characteristics, such as half-life, effector functions and stability, which affect the suitability for manufacturability and therapeutic purposes [22]. IgG1 is the most commonly used subclass due to characteristics such as high affinity to Fc receptors and long plasma half-life [21]. Of note, a recent review suggests to reconsider the therapeutic potential of engineered IgG3 due to enhanced effector functions and better antigen accessibility [23]. In addition, the different subclasses of IgG can be further subdivided into distinct allotypes [24, 25]. The Kabat numbering is commonly used to describe positions of amino acid residues [26]. It was the first standardized numbering system and facilitated the comparison of different antibodies.

IgGs carry a conserved *N*-glycosylation site in the heavy chain (CH2 domain) at Asn297 [27]. Since IgGs exhibit two glycosylated heavy chains, heterogeneous glycan pairings are commonly present (**Figure 1.2**). Due to the heterogeneity of glycan features and the different combination possibilities of mAb glycan pairings, the functional understanding of glycosylation is a major analytical challenge [28]. Further, non-canonical Fab *N*-glycosylation may be present in the variable domains [29]. Of note, *O*-glycans have also been reported for elongated therapeutic mAbs [30]. For the conserved Fc *N*-glycosylation, a dedicated nomenclature (GxF) is commonly used for describing complex-type diantennary glycans of IgGs. Gx indicates the number (x) of terminal galactoses and F the presence of core fucose. In addition, GxF-N (monoantennary), GxFN (bisection) and Mx (high mannose, x = number of mannoses) are differentiated [12].

1.1.3. FUNCTIONAL RELEVANCE OF IMMUNOGLOBULIN G GLYCOSYLATION

Glycosylation features of IgG Fc *N*-glycans are known to impact effector functions [12, 29]. For example, the interaction of IgG and the Fc gamma receptor III (FcγRIII) is known to be highly glycosylation dependent, which is attributed to unique glycan-glycan interactions [31]. Two different types of FcγRIII are expressed on natural killer cells or neutrophils. Both types initiate key immunological processes such as antibody-dependent cellular cytotoxicity or antibody-dependent cellular phagocytosis [32]. Hence, the effector functions of a therapeutic mAb are greatly affected by its glycosylation features. Fucosylation of IgG Fc glycans is known to tremendously (up to 50 fold) decrease the binding affinity towards FcγRIII [33]. Other glycosylation features

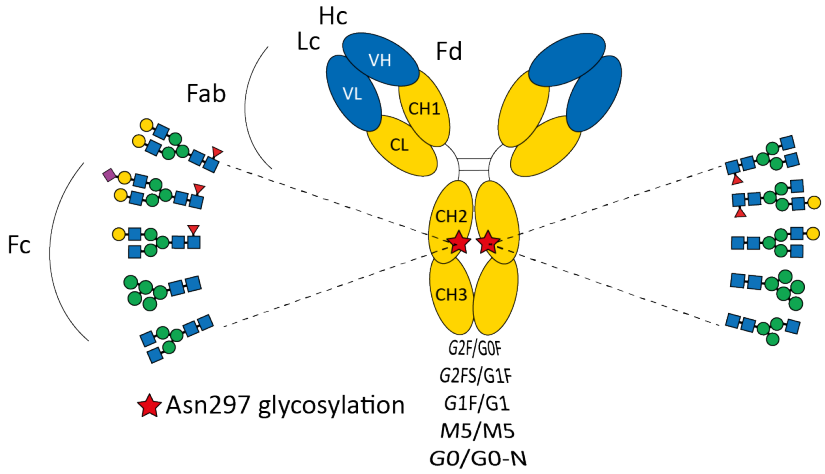

Figure 1.2: Scheme of IgG1-based mAb and commonly observed Asn297 glycans.

such as galactosylation or bisection also influence the binding affinity, although to a smaller extend [8, 34, 35]. Glycoengineering approaches to decrease fucosylation have been successfully proven to enhance the clinical efficacy of mAbs such as Gazyva® [29, 36, 37]. Besides effector functions, the pharmacokinetics are also affected by Fc glycosylation. For example, hybrid and high mannose glycoforms are cleared faster from circulation [38].

1.1.4. Other post translational modifications of proteins

Besides glycosylation, additional post translational modifications (PTMs) occur naturally, or they are introduced accidentally during manufacturing, purification or storage of therapeutic proteins (**Figure 1.3**) [39]. Common PTMs are C-terminal Lys/Arg clipping-variants, deamidation (Asn, Gln) or oxidation (Met, Trp, Cys). Certain PTMs are known to impact the function of therapeutic proteins and hence are defined as critical quality attributes (CQAs) [40]. For example, Asn deamidation of mAbs has been reported to affect antigen binding or Fc receptor binding [41–43]. Fc oxidation decreases half-life, owing to decreased neonatal Fc receptor binding and scavenging [44, 45]. Co-occurring PTMs on therapeutic proteins cause a high degree of complexity. Glycosylation is the most complex PTM and hence a major contributor to proteoform heterogeneity. For complex glycoproteins such as EPO, several hundred proteoforms have been reported [46]. The FDA estimated the presence of 108 theoretical proteoforms for a therapeutic mAb [47]. However, it is critical to understand which of these proteoforms are abundant enough to cause a significant biological/clinical effect. Hence, the analytical characterization of PTMs is crucial and the elucidation of structure-function relationships is of high importance to assess the criticality of a PTM [48].

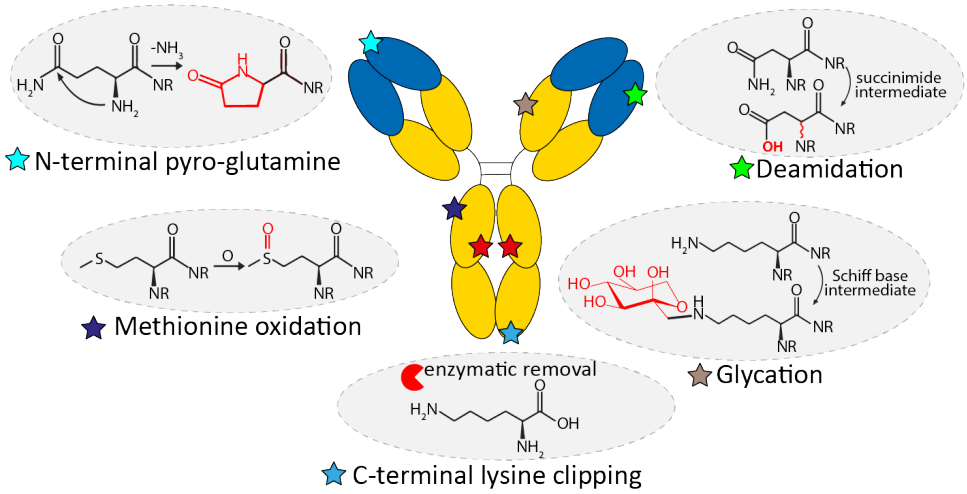

Figure 1.3: Visualization of different post-translational modifications commonly present in therapeutic proteins.

1.2. Mass spectrometric analysis of therapeutic glycoproteins

Tremendous technological improvements in mass spectrometry (MS) instrumentation have been made in the last decades, which redefined the way (therapeutic) glycoproteins are analyzed [49–53]. The gold standard for glycosylation analysis was based on hydrophilic interaction liquid chromatography of released glycans with fluorescence detection [54]. However, technological advances allow for a protein-specific glycosylation analysis with MS detection. MS is a versatile technique, which is used extensively for the characterization of therapeutic glycoproteins in academia and industry [39, 55, 56]. In MS, the mass to charge ratio (m/z) of an ionized analyte is determined in the gas phase. The two most common ionization techniques for large biomolecules, such as proteins, are matrix assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) [57, 58]. Both techniques allow for soft ionization of biomolecules, i.e. non-destructive ion formation. Several analytical MS-based approaches have been developed for (therapeutic) proteins and PTMs, namely bottom-up, middle-up and intact mass analysis (**Figure 1.4**). For a comprehensive analytical characterization of therapeutic proteins, the integration of multi-level analyses is pivotal [59–61].

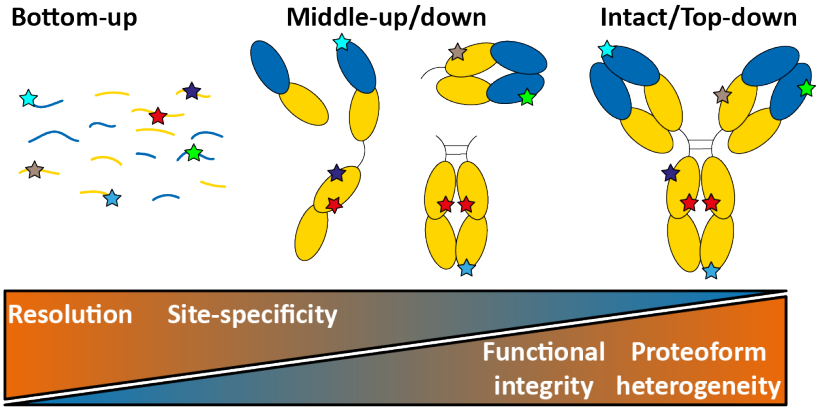

Figure 1.4: Schematic representation of different approaches for the MS analysis of mAb proteoforms.

1.2.1. IONIZATION TECHNIQUES

MALDI was introduced by Karas and Hillenkamp [62, 63]. Several advantages, such as short analysis time, low spectrum complexity and straightforward data interpretation, make MALDI highly suitable for high-throughput applications. In MALDI, the analyte is first co-crystallized with a MALDI matrix on a target plate. Small aromatic compounds, which show absorption in the ultraviolet or infrared wavelengths, are usually used as MALDI matrices. A laser pulse irradiation under vacuum causes a rapid heating of the crystals. Of note, the exact ionization mechanism of MALDI is not fully grasped. Ion formation is suggested either via a proton transfer mechanism in the solid phase prior to desorption or in the gas phase upon ablation [64, 65]. Predominantly singly charged ions are produced by MALDI, which simplifies the data interpretation. Most crucial for successful MALDI experiments is the selection of an appropriate matrix, which depends on the analyte of interest [66].

ESI of proteins was first described by Fenn et al. [58]. Separation techniques may be hyphenated to MS via ESI, which is highly advantageous when analyzing complex mixtures. In ESI, the liquid flow is dispersed into a spray under high voltage and highly charged droplets are produced. Solvent molecules are evaporated, and ions are formed via desorption of charged analytes from the surface of the droplets or after their complete desolvation. In contrast to MALDI, ions produced by ESI are multiply charged. This increases the potential mass range of analytes and MS resolution. Furthermore, ESI allows for analyzing proteins or protein complexes retaining a native-like structure in the gas phase [67].

1.2.2. BOTTOM-UP ANALYSIS

In bottom-up approaches, proteins are analyzed at the peptide level. The advantages are the high resolution and the site-specific information obtained on modifications. In bottom-up workflows, proteins are usually reduced and alkylated followed by protease cleavage. Trypsin is the gold standard protease. However, other specific proteases exist and are useful alternatives to obtain complementary information [68]. Peptides are then separated, mostly using reversed-phase (RP) liquid chromatography (LC) prior to ESI-MS. For identification purposes, peptides are fragmented e.g. via gas-phase collisions or radical transfer reactions to assign sequences and potential PTMs. In the biopharmaceutical industry, bottom-up methods are popular for multiple-attribute monitoring, meaning the simultaneous assessment of different PTMs. This is related to the higher assignment confidence and specificity achieved by bottom-up approaches compared to middle-up or intact analysis. However, the actual information of proteoforms is lost and artificial modifications may be introduced during the long sample preparation [69]. Bottom-up methods are very data-rich and efficient data processing remains a bottleneck, in particular for glycopeptides.

1.2.3. INTACT ANALYSIS

Intact analysis of proteins is performed without prior chemical or enzymatic sample preparation. In contrast to bottom-up approaches, information about the combination of PTMs on one molecule is revealed, e.g. pairing of glycans on a mAb. Further, better insights on the actual proteoform heterogeneity are achieved since the distribution of PTMs, i.e. glycans, does not follow a random statistical distribution. However, intact protein analysis lacks sensitivity and has a high degree of proteoform assignment ambiguities, because intact mass profiles may be very complex. Particularly, glycoproteins with multiple glycosylation sites, such as EPO, suffer from unresolved isobaric and isomeric PTMs [59].

Intact protein analysis by MS is greatly facilitated by online hyphenation to separation techniques such as size-exclusion chromatography (SEC) or ion exchange chromatography (IEC) [70, 71]. Alternatively, Fc receptor affinity chromatography (AC) is a recently introduced strategy for functional separation of mAb proteoforms [44, 72]. The commercialization of Fc receptor affinity columns as analytical tool is currently emerging. AC allows to decomplexify intact glycoprotein mixtures. However, hyphenation of functional AC and MS is yet unexplored in literature for mAb analysis.

1.2.4. MIDDLE-UP ANALYSIS

As alternative to intact protein analysis, middle-up strategies are often applied for the analysis of antibodies, since they reduce the proteoform heterogeneity and increase the MS resolution and sensitivity. Workflows consist of hinge-specific proteolysis and/or simple reduction of disulfide bonds. Due to vastly different ionization efficiencies of the subunits, middle-up experiments are commonly performed in combination with a

separation dimension. In recent years, hinge-specific proteases have been commercialized and have facilitated the subunit analysis of therapeutic IgGs [73–75]. IgG hinge-proteases are very specific and require only a short sample preparation. In addition, some hinge protease can cleave under native conditions, which allows to analyze the distinct function of Fc and Fab fragments. Middle-up experiments complement the information obtained by bottom-up or intact protein analysis and are a good compromise of structural integrity and site-specificity. Hence, this approach is very helpful for assigning PTMs to a specific subunit.

1.3. SCOPE

The analytical characterization of therapeutic glycoproteins is crucial to assess their safety and efficacy. However, the complexity of glycosylation is a major challenge for analytical methods. The aim of the work in this thesis was to develop new mass spectrometry-based techniques for the characterization of therapeutic glycoproteins. The developed methods address current issues in glycoproteomic data-processing and sample preparation and show the feasibility of MALDI for intact glycoform profiling. They also demonstrate the power of intact- and middle-up mAb Fc receptor affinity chromatography - mass spectrometry for proteoform-resolved structure-function studies.

In **Chapter 1** the importance of therapeutic glycoproteins is emphasized. PTMs, with focus on glycosylation, are described with respect to their impact on proteoform heterogeneity and structure-function relationships. Further, the most common ionization techniques for protein analysis, MALDI and ESI are briefly introduced. Finally, different MS-based analytical strategies for the characterization of mAbs are elaborated.

Bottom-up approaches using RP LC-MS/MS are commonly used for glycoproteomic experiments. These approaches are very powerful for comprehensive glycosylation characterization. However, the vast complexity of data remains a major bottleneck in glycoproteomic workflows. In **Chapter 2,** bioinformatic tools are combined into a semi-automated workflow for glycoproteomics data analysis. Glycopeptide coverage and the reliability of relative quantification could be increased.

Protease selection is a crucial step for successful glycoproteomic experiments. One important selection criterion for a protease is that it generates singly glycosylated glycopeptides. In **Chapter 3** an alternative cysteine alkylation strategy, aminoethylation, is applied for the first time for the bottom-up analysis of EPO. Aminoethylation allows tryptic cleavage at newly introduced cleavage positions between two glycosylation sites, which could be previously only achieved by less specific proteases.

Intact protein analysis has the unique ability to reveal the combination of glycoforms of multiply glycosylated proteins. MALDI MS is usually not performed for comprehensive intact glycoform profiling, owing to the low glycoform resolution and low stability of glycosylation features, i.e. sialic acids. **Chapter 4** demonstrates the feasibility of glycoform-resolved intact analysis of EPO by MALDI Fourier-transform ion cyclotron resonance MS. The method shows a high stability of sialic acid moieties and facilitates the data analysis as no charge deconvolution is required.

Intact mass analysis of mAbs is commonly done by direct infusion MS or separation-based techniques such as RP-HPLC MS. However, glycoforms are usually not separated in this manner and no functional information is derived. In **Chapter 5**, a method for functional FcγRIIIa AC with online MS hyphenation is presented. New insights into the FcγRIIIa affinity of mAb glycoform pairings are derived from this analytical setup.

The glycoform assignments of intact mAbs are highly ambiguous, if additional glycosylation is present in the Fab. To expand the applicability of the developed method from Chapter 5, different IgG hinge proteases are evaluated in **Chapter 6** for an FcγRIIIa AC-MS method of Fc-moieties. The advantages with respect to proteoform resolution and assignment ambiguities, are demonstrated for Fab-glycosylated cetuximab.

In **Chapter 7**, an AC-MS method is described for FcγRIIIb mAb interactions. This is the first report on an FcγRIIIb AC, demonstrating the suitability of AC to study low-affinity interactions.

A discussion about the challenges and outlooks of MS data analysis for glycoproteins and the potential of functional AC-MS for biopharmaceutical research and development is provided in **Chapter 8**.

2

SEMIAUTOMATED GLYCOPROTEOMICS DATA ANALYSIS WORKFLOW FOR MAXIMIZED GLYCOPEPTIDE IDENTIFICATION AND RELIABLE QUANTIFICATION

Steffen Lippold, [1] Arnoud H. de Ru, [1] Jan Nouta,[1] Peter A. van Veelen,[1] Magnus Palmblad,[1] Manfred Wuhrer[1] and Noortje de Haan[1]

[1]Center for Proteomics and Metabolomics, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, Netherlands

Reprinted and adapted from Beilstein J. Org. Chem. 2020, 16, 3038–3051
DOI: 10.3762/bjoc.16.253

Glycoproteomic data are often very complex, reflecting the high structural diversity of peptide and glycan portions. The use of glycopeptide-centered glycoproteomics by mass spectrometry is rapidly evolving in many research areas, leading to a demand in reliable data analysis tools. In recent years, several bioinformatic tools were developed to facilitate and improve both the identification and quantification of glycopeptides. Here, a selection of these tools was combined and evaluated with the aim of establishing a robust glycopeptide detection and quantification workflow targeting enriched glycoproteins. For this purpose, a tryptic digest from affinity-purified immunoglobulins G and A was analyzed on a nano-reversed-phase liquid chromatography–tandem mass spectrometry platform with a high-resolution mass analyzer and higher-energy collisional dissociation fragmentation. Initial glycopeptide identification based on MS/MS data was aided by the Byonic software. Additional MS1-based glycopeptide identification relying on accurate mass and retention time differences using GlycopeptideGraphMS considerably expanded the set of confidently annotated glycopeptides. For glycopeptide quantification, the performance of LaCyTools was compared to Skyline, and GlycopeptideGraphMS. All quantification packages resulted in comparable glycosylation profiles but featured differences in terms of robustness and data quality control. Partial cysteine oxidation was identified as an unexpectedly abundant peptide modification and impaired the automated processing of several IgA glycopeptides. Finally, this study presents a semiautomated workflow for reliable glycoproteomic data analysis by the combination of software packages for MS/MS- and MS1-based glycopeptide identification as well as the integration of analyte quality control and quantification.

2.1. INTRODUCTION

PROTEIN glycosylation mainly occurs in the form of *N*- and *O*-glycosylation. *N*-Glycans are attached to Asn within an amino acid consensus sequence (Asn-Xxx-Ser/Thr, Xxx ≠ Pro) and *O*-glycans are attached to Ser or Thr. Glycan compositions can range from monosaccharides (e.g., Tn antigen for *O*-glycans [76]) to large polysaccharides (e.g., *N*-glycans of recombinant human erythropoietin [77]). The most common building blocks of human protein glycans are hexoses (glucose, galactose, and mannose, Hex/H, 162.0528 Da), *N*-acetylhexosamines (*N*-acetylglucosamine or *N*-acetylgalactosamine, HexNAc/N, 203.0794 Da), fucose (Fuc/F, 145.0579 Da), and sialic acid (*N*-acetylneuraminic acid, NeuAc/S, 291.0954 Da). The combinatorial possibilities of these building blocks and the variety of structural features, such as the linkage position and anomeric configuration, make protein glycosylation a highly complex posttranslational modification (PTM).

Glycoproteomics has become important for many life science disciplines, in particular for biomedical and biopharmaceutical research [78–80]. Glycopeptide-centered glycoproteomics aims at the characterization of macroheterogeneity and microheterogeneity of protein glycosylation [13]. Reversed-phase liquid chromatography coupled to high-resolution tandem mass spectrometry (RPLC–MS/MS) is a standard analytical method in the field of glycoproteomics [81]. The separation of glycopeptides in RPLC is mainly driven by the peptide portions. Thus, information on different proteins and glycosylation sites appears in the form of glycopeptide clusters. Next to the peptide portion, glycosylation features, such as sialic acids, can strongly influence the retention time [82]. Advances in MS technologies tremendously enhanced the detection and informative fragmentation of glycopeptides in the past years [83]. The large amount of highly complex data acquired using these technologies shifted the major bottleneck in glycopeptide analysis to the data processing steps. Next to the high complexity of glycosylation itself, data analysis is further complicated by interfering background signals from biological matrices and isomeric and near-isobaric ambiguities resulting from combinations of monosaccharides, adducts, amino acids, and amino acid modifications [84, 85].

Efforts have been made in recent years in the development of bioinformatic tools to facilitate and automate data processing in glycopeptide-centered glycoproteomics [86]. Several reports have reviewed the functionalities and application areas of data analysis tools in the field of glycoproteomics [81, 83, 86, 87]. MS/MS-based scoring software tools such as Byonic [88] are frequently used for glycopeptide identification [86]. Recently, software tools were developed that are based on the retention time (RT) characteristics and accurate mass differences of glycopeptide MS1 signals in RPLC–MS [84, 89]. These tools detect inaccuracies of MS/MS assignments based on the RT and increase the number of identified glycopeptide compositions while keeping the false positive assignments low. Other reports performed glycopeptide identification using summed MS1 spectra of previously defined elution clusters [90]. This approach is applicable when the identity and elution behavior of the glycopeptides of interest is

known and is aided by quality criteria such as mass accuracy and isotopic pattern matching. Furthermore, such approaches allow quantification in a high-throughput manner, which is advantageous e.g., in clinal cohort analysis [90–92].

Here, we present a workflow for the reliable and efficient analysis of glycopeptides from enriched glycoproteins. We performed a thorough evaluation of the software tools and workflows used in our laboratory for the identification and quantification of glycopeptides. For this, a sample containing immunoglobulins G and A (IgG and IgA), simultaneously captured from human plasma, was chosen. This sample showed a considerable level of complexity due to the presence of multiple glycoproteins of interest and cocaptured (glyco)proteins from the plasma. The tools included Byonic, GlycopeptideGraphMS, Skyline, and LaCyTools.

2.2. Results and Discussion

2.2.1. Glycoproteomics data analysis workflow

Affinity-copurified IgG/IgA from human plasma was chosen as a sample to demonstrate the integration of tools for the semiautomated glycoproteomic data analysis (**Figure 2.1**). The three main parts of this workflow cover glycopeptide identification (Byonic, GlycopeptideGraphMS), curation, and quantification (LaCyTools).

In the first step, Byonic was used for automated MS/MS-based (glyco)peptide identification. This initial step is crucial to validate the presence of glycopeptides and the assignment of the peptide portions. Next, the number of identified glycopeptides was maximized by performing an open search based on MS1 information (mass and RT) in GlycopeptideGraphMS. A preprocessing step in OpenMS was performed as described for the original GlycopeptideGraphMS workflow [89], including deisotoping and decharging of all features. The outcome of GlycopeptideGraphMS is a list with glycopeptide clusters (defined as LC–MS features (nodes) that are connected by Δmass and ΔRT within the provided limits for glycopeptides), for which at least one node should be confidently assigned by MS/MS to identify all glycopeptides in a cluster. The clusters are also presented in interactive graphs, which assist in the identification of false-positive connections (unlikely mass/RT shifts) and unexpected glycopeptide clusters (e.g., missed cleaved products and peptide or glycan modifications). This information can be used in an iterative manner to adjust the search space for Byonic. Study-specific search criteria are listed in the Experimental section and a detailed manual for the use of GlycopeptideGraphMS can be found elsewhere [89]. Of note, separate LC–MS runs with exclusively MS1 information were acquired in order to maximize the MS1-based identification and to ensure the highest possible data quality for the quantification purposes.

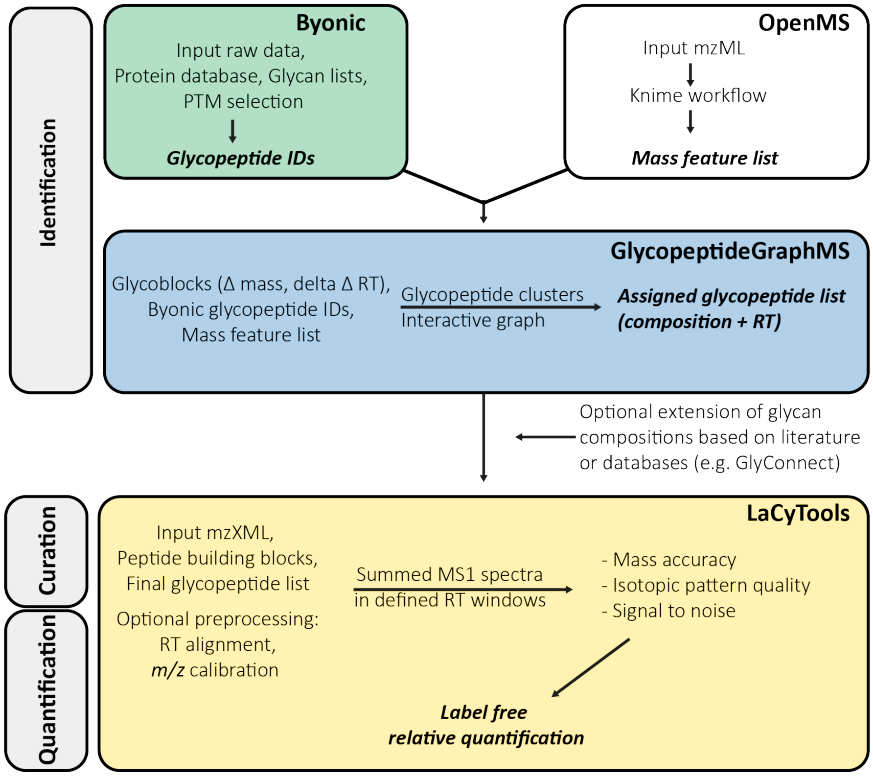

Figure 2.1: Integration of automated glycopeptide identification by Byonic and GlycopeptideGraphMS (aided by OpenMS) and subsequent analyte quality control and quantification by LaCyTools.

Upon glycopeptide identification, the list of glycopeptides generated by GlycopeptideGraphMS was transformed to the input format required for targeted curation and quantification in LaCyTools [90]. A python script was developed to facilitate this step (**Supporting Information File 3**). LaCyTools was chosen because it is open-source, can be applied for a large number of samples (thousands of samples in one study have been reported [93]), and allows data curation and quantification. Importantly, LaCyTools requires RT clusters to be defined in which MS1 spectra can be summed and further processed, which is facilitated by the GlycopeptideGraphMS output. The analyte list may be extended by including glycan compositions (e.g., from the literature or databases such as GlyConnect [94]) within appropriate RT clusters (e.g., the same peptide portion and number of sialic acids). Furthermore, the user has the option to perform preprocessing steps, such as *m/z* calibration and RT alignment. For data curation, summed MS1 spectra were subjected to quality control based on user-defined cut-offs for mass accuracy, isotopic pattern matching, and the signal-to-noise ratio of an analyte. Finally, the integrated areas of all charge states

passing the quality criteria were summed for each glycopeptide composition, the area was corrected for missing isotopes, and total area normalization was performed for label-free relative quantification. Study-specific parameters for the use of LaCyTools are provided in the Experimental section and further explanation on the use of this tool can be found elsewhere [95].

2.2.2. GLYCOPEPTIDE IDENTIFICATION

AUTOMATED MS/MS-BASED GLYCOPEPTIDE IDENTIFICATION BY BYONIC

The automated and score-based MS/MS glycopeptide identification using Byonic resulted in the confident assignment of ten IgG/IgA *N*-glycopeptide clusters of interest (**Table 2.1** and **Figures S1–S10**, **Supporting Information File 2**).

Table 2.1: Automated MS/MS-based identification of IgG/IgA glycosylation sites by Byonic. For each glycopeptide moiety, a representative glycoform is shown (see **Figures S1–S10**, **Supporting Information File 2** for the corresponding MS/MS spectra).

Protein	Glycopeptide	Glycosylation site[a]	Cluster	Mass error (ppm)	Score	Scan time (min)
IgG1	R.EEQYN[+H5N4F1]STYR.V	Asn297	IgG1	0.7	589	14.4
IgG2/3	R.EEQFN[+H3N4F1]STFR.V	Asn297	IgG2/3	0	693	18.5
IgG4	R.EEQFN[+H3N4F1]STYR.V	Asn297	IgG4	1.1	401	15.8
IgA1/2	R.LSLHRPALEDLLLGSEAN[+H5N4S1]LTC[+57]TLTGLR.D	Asn263	LSL	0.9	839	40.2
	R.LAGKPTHVN[+H5N5F1S2]VSVVM[+16]AEVDGTC[+57]Y.-[b]	Asn459	LAGY	0.4	601	25.5
	R.LAGKPTHVN[+H5N5F1S2]VSVVM[+16]AEVDGTC[+57].-[b]	Asn459	LAGC	2.9	649	25.9
IgA2	K.TPLTAN[+H5N4F1S1]ITK.S	Asn337	TPL	-1.2	728	19.1
	K.HYTN[+H5N5F1S1]SSQDVTVPC[+57]R.V	Asn211	HYT	1.3	194	15.6
JC	R.EN[+H5N4S2]ISDPTSPLR.T	Asn49	ENI	0.1	565	22.2
	R.IIVPLNNREN[+H5N4F1S1]ISDPTSPLR.T	Asn49	IIV	1.2	271	28.0

[a] Numbering according to [92]. [b] C-terminal peptide of the heavy chain, no C-terminal tryptic cleavage.

Assigned glycopeptides from copurified human plasma proteins other than IgG and IgA were not considered for further data processing (e.g., fibrinogen, alpha-1-antitrypsin, or clusterin, see **Table S1A–E**, **Supporting Information File 1**). Missed-cleavage variants were assigned for IgG1, IgG2/3, and IgA1/2 (Asn263) but not further considered because of their low abundance. For the IgA joining chain (JC), the elongated peptide with a missed cleavage was included for further data processing as the cleavage efficiency was previously determined to be glycoform dependent [92, 96]. For the assignment of tryptic *N*-glycopeptides to specific proteins, ambiguities exist for one peptide moiety that could be assigned to either IgG2 or IgG3 and three moieties that were shared between IgA1 and IgA2 (**Table 2.1**) [78]. These ambiguities were not resolved using the proposed workflow. However, the presence of protein-specific (non)glycopeptides may indicate differences in the abundance of the individual proteins. For addressing these ambiguities, a more selective sample preparation is required, for example, using different enrichment strategies or proteases [97]. Interestingly, an additional allotype of the main IgG3 glycosylation site (EEQYNSTFR) was assigned in four out of five technical replicates by Byonic. This IgG3 glycopeptide is an isomer of the tryptic IgG4 glycopeptide (EEQFNSTYR). However, upon manual inspection of the data, only one scan of the assigned MS/MS spectra within all five

technical replicates covered the relevant amino acids (position of Phe and Tyr), allowing an unambiguous discrimination between IgG3 or IgG4 (score 281, **Figure S11**, **Supporting Information File 2**). The IgA2 HYT glycopeptides had the lowest scores (max. 194) compared to the other glycosylation sites. It was detected in four out of five technical replicates and only with a maximum of one glycan composition. The low intensity of these glycopeptide signals resulted in a decreased likelihood for MS/MS selection. Of note, the IgA2 HYT glycopeptide covers a sequence stretch homologous to the hinge region of IgA1, carrying *O*-glycans. In a previous study the IgA1 peptide has been referred to as the HYT glycopeptide cluster as well [91]. The C-terminal IgA1/2 glycopeptides (LAGC/Y) were found mainly with methionine oxidation. Unoxidized peptide moieties were also assigned but with low scores (below 50). The manual check of the data revealed that in some cases, the selection of the wrong monoisotopic mass in Byonic led to misassignments of near-isobaric compositions, e.g., TPL H5N5F3 (3+, *m/z* 1074.8020, false) instead of TPL H5N5F1S1 (3+, *m/z* 1074.4619, correct). Other theoretical possible, but less common, tryptic IgG3, IgA1, and IgA2 glycopeptides were not detected [78, 91]. One of the reported common miscleaved IgA2 *N*-glycopeptides (SESGQNVTAR) is likely to elute prior to MS acquisition as described previously for the applied gradient [91]. For the expected IgA1 *O*-glycopeptide cluster, the Byonic search failed to score any hits when performed as described previously [91]. Of note, the tryptic *O*-glycopeptide cluster could be detected upon manual inspection, albeit with low intensity (**Figures S12** and **S13**, **Supporting Information File 2**). The reason for this was further investigated based on the GlycopeptideGraphMS results and is discussed in the section on automated MS1- and RT-based glycopeptide identification by GlycopeptideGraphMS.

In MS/MS scoring approaches such as Byonic, the definition of a threshold for the automated assignment of glycopeptides is generally a challenge as the scores depend largely on the fragmentation method, the peptide characteristics (e.g., peptide length or additional modifications), the glycome, and the sample matrix [85]. A recent study by us applied a threshold score of 200 for the IgG/IgA glycopeptides from human serum, aiming to find a balance between the exclusion of false positives while preventing false negatives [91]. Sensitive glycopeptide assignments relying only on oxonium ions and precursor mass, using a score above 30 were also described recently [89]. A suitable cut-off score should always be carefully evaluated for each (glyco)peptide moiety with respect to the glycoform coverage and accuracy [85].

Byonic identified the relevant *N*-glycosylation sites of IgG/IgA in all five technical replicates with the exception of the low-abundant IgA2 HYT glycopeptide. Further results and discussion of the accuracy and coverage of the investigated glycopeptides of interest are presented in the following section. Software tools for automated MS/MS-based assignments such as Byonic are highly useful in glycoproteomic data processing workflows. Other, noncommercial, automated MS/MS-based software tools for glycopeptide identification were recently reviewed [86] and have the potential to substitute Byonic in similar workflows as described here. However, these tools were not evaluated in the current study.

The glycopeptide identification was further extended by an open MS1 search based on mass and RT differences using GlycopeptideGraphMS [89]. RT clusters for all MS/MS assigned IgG/IgA glycopeptides were found using this tool (**Figure 2.2**). Of note, GlycopeptideGraphMS relies on the MS/MS assignment of at least one glycopeptide per RT cluster (be it automated or manual). The GlycopeptideGraphMS cluster with the highest number of connections contained the expected masses of the IgA1 *O*-glycopeptides, which were not assigned in the Byonic search (**Figure S12**, **Supporting Information File 2**). In line with the Byonic search, several other RT clusters of missed cleaved products or glycopeptides from other plasma proteins were present (data not shown).

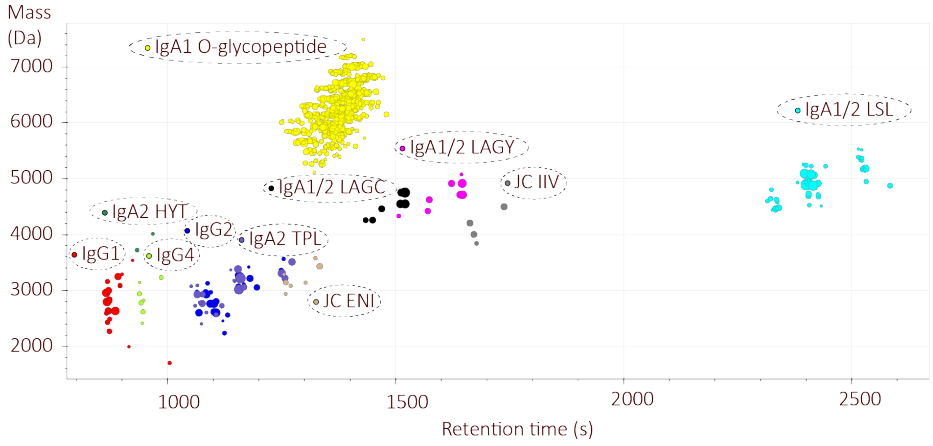

Figure 2.2: Representative IgG and IgA glycopeptide clusters detected by GlycopeptideGraphMS.

Additional clusters with a +27.9949 Da (formylation) mass shift and an increased RT were observed for most of the IgG and IgA glycopeptides (see **Figure S14**, **Supporting Information File 2** for representative IgG glycopeptide examples). The formylation was conveniently assigned to the glycan part (**Figure S15**, **Supporting Information File 2**) but may occur at the peptide portion as well [98–100]. Formylation is likely introduced by the exposure of the tryptic peptides to formic acid during the acid precipitation of sodium deoxycholate in the final step of the sample preparation and during subsequent storage [98]. Within the glycopeptide clusters of interest, Cys oxidation (+15.9949 Da) was assigned as an unexpected modification in all Cys-containing glycopeptides (five out of 11) at a high relative abundance (65.4–77.2%) and confirmed upon manual inspection of the MS/MS data (**Figures S13** and **S16–S18**, **Supporting Information File 2**). The y- and Y-fragment ions of (glyco)peptides with Cys oxidation showed a

characteristic neutral loss of 107.0041 Da (C2H5O2NS), as reported for singly oxidized carbamidomethylated Cys through an elimination reaction in the gas phase (**Figures S13** and **S16–S18**, **Supporting Information File 2**) [101]. Peptides with Cys oxidation had a similar elution behavior as the unoxidized isomeric counterparts (with an additional hexose instead of a fucose unit), leading to a high degree of ambiguous, albeit often illogical compositions (e.g.. for the LSL cluster, **Figure S16**, **Supporting Information File 2** and **Table S2A–E**, **Supporting Information File 1**) and false-positive assignments (e.g., the LAGC cluster, **Figure S17**, **Supporting Information File 2**) in GlycopeptideGraphMS. In line with these findings, the high number of illogical compositions and false-positive assignments of the IgA1 *O*-glycopeptide (three Cys residues) were due to modification variants on the Cys residue (**Figures S12** and **S13**, **Supporting Information File 2**). In general, the assignment based on RT differences and MS1 information (manual or automated) had a highly increased uncertainty for the glycopeptides with partial Cys oxidation, and MS/MS was essential for confident identification in these cases. Of note, false-positive assignments related to Cys oxidation were also observed in the automated Byonic search upon manually reevaluation. For example, the LAGC glycopeptide composition H6N5S2 had a maximum score of 282, with no coverage of y-ions (**Figure S17**, **Supporting Information File 2**). This was due to the presence of the oxidized Cys residue at the C-terminus for which characteristic y- and Y-ions could be manually assigned in this scan. These findings substantiate that the scores in automated MS/MS searches may be still relatively high for false-positive assignments. Defining the appropriate search space with prior knowledge on relevant modifications and neutral losses is crucial to increase the identification accuracy for (glyco)peptides with unexpected modifications, such as Cys oxidation. The oxidation of Cys can appear biologically in the sample or artificially during/upon sample preparation [101, 102]. In general, Met modifications are known for causing ambiguities in glycoproteomics due to partial oxidation, particularly in combination with carbamidomethylation [85, 103]. To our knowledge, no study has previously reported on partial Cys oxidation as a confounder in glycoproteomics. As peptides containing the Cys oxidation had a higher abundance than the unoxidized counterparts, it is stressed that this modification should be carefully checked in Cys-containing glycopeptides as in the investigated sample, it had major implications on the IgA glycoprofiling accuracy. Further elaboration of the Cys-containing peptides, including modifications and correct glycan composition identifications, were considered beyond the scope of this study due to the largely increased complexity. Hence, the applicability of the proposed glycoproteomic data analysis workflow was demonstrated on a subset of six *N*-glycopeptide clusters, namely IgG1, IgG2/3, IgG4, JC (ENI, IIV), and IgA2 (TPL).

For the six glycopeptide clusters of interest, the presence of 262 theoretical glycopeptides (based on the internal IgG/IgA glycan reference list [91] and Glyconnect entries for these peptides [94], **Table S3**, **Supporting Information File 1**) was manually evaluated in Skyline, and the presence of 83 glycopeptides in the used data was confirmed (**Table S4**, **Supporting Information File 1**). In total, 82 correct glycopeptide compositions were identified using GlycopeptideGraphMS with MS/MS validation,

whereas the Byonic-only search resulted in 35 compositions (**Table S4**, **Supporting Information File 1**). Of note, four glycan compositions (H2N3F1, H2N4F1, H5N3F1S1, H5N5F2S1) were not included in the *N*-glycan search list of Byonic, and hence not included for the calculation of its glycopeptide coverage. Those glycans were only present in low abundance on the glycopeptides, and often no MS/MS spectrum was present (**Figure 2.3**). However, it highlights the importance of a complete glycan composition list for a database-based identification of glycopeptides, something that is less critical in MS1-based RT and accurate-mass-difference searches.

In the GlycopeptideGraphMS search, nine compositions were detected that were not within the internal IgG/IgA glycan reference list [91] or had an entry in Glyconnect for these peptides (**Table S4**, **Supporting Information File 1**) [94]. These analytes were present at very low relative abundances (<1%). The IgA2 TPL peptide showed the highest number (five) of additional compositions (H6N5F1, H5N5, H5N4F2S1, H5N5F2S1, and H6N4F1S1). For all glycoforms identified by GlycopeptideGraphMS, only one composition (TPL, H4N4F2S1) was determined as false-positive as no MS1 signals could be found for this analyte in the raw data. Of note, one TPL glycoform (H5N5S1) was detected with a low abundance in three out of five technical replicates but was excluded from the identification and further processing due to the presence of isobaric MS signals in the raw data. On the other hand, only one glycopeptide (IgG2/3 H3N3F1) was assigned manually, without being identified by GlycopeptideGraphMS as the correct mass and RT combination was not in the deconvoluted mass list. These false-positive and false-negative results are artifacts of the feature recognition, deconvolution, and deisotoping in OpenMS [89] prior to the GlycopeptideGraphMS analysis. Furthermore, the preprocessing steps caused some glycopeptides to be detected at multiple RT values (**Figure 2.3**), whereas the raw data showed only a single chromatographic peak. The applied OpenMS workflow has a reported accuracy of 91% for detecting the correct monoisotopic peak of a feature, and this workflow was not further optimized in the current study [89].

With respect to the consistency of the glycopeptide identification in technical replicates, the MS1-based identification (GlycopeptideGraphMS) supported by MS/MS data showed a better performance than MS/MS identification alone (47 vs 16 glycopeptides detected in all replicates, respectively, **Figure S19**, **Supporting Information File 2**). Both automated identification approaches showed variations within the data of the technical replicates, and the glycopeptide coverage was maximized by combining all measurements. For the MS1-based assignment, the variation between replicates was found in the minor glycan species, which were on the borderline of the limit of detection. Further, the stochastic nature of MS/MS selection is a known factor, which may cause variability in MS/MS-based assignments [89].

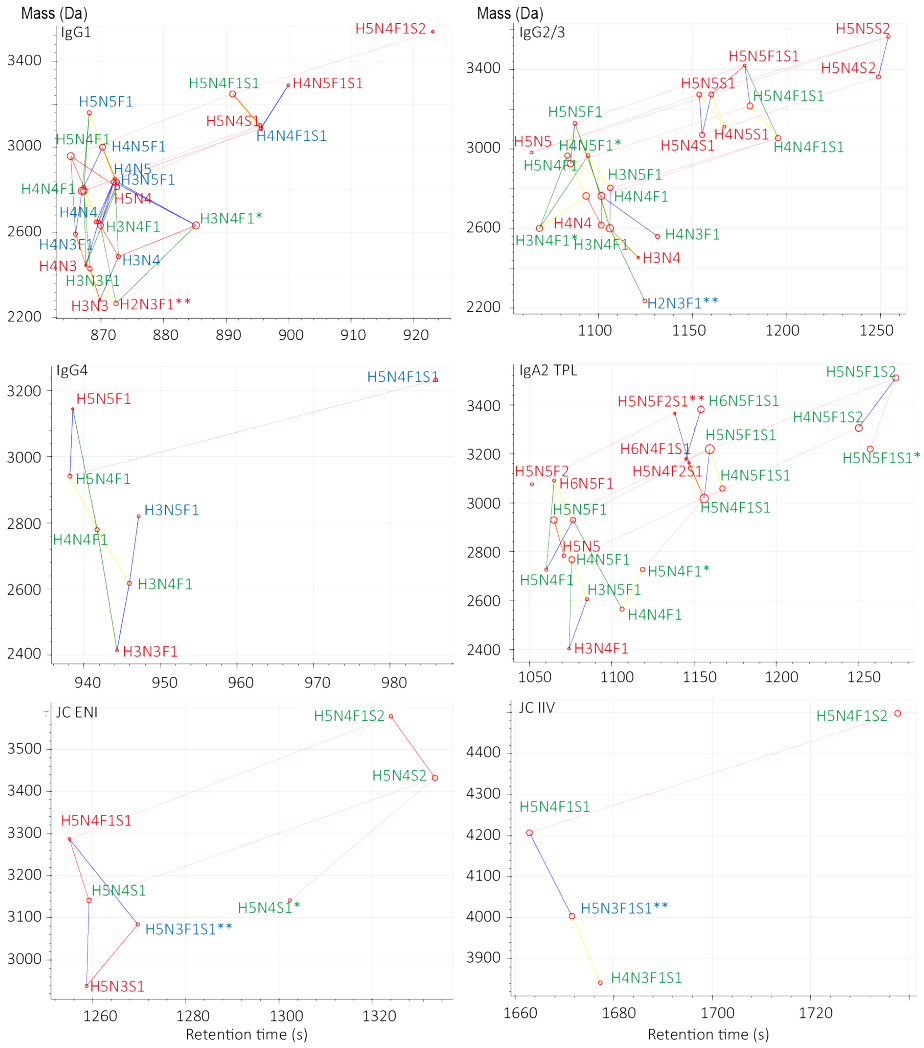

Figure 2.3: Representative GlycopeptideGraphMS output for peptides of interest. Assigned compositions were identified using MS/MS data via Byonic (green) or manual assignment (blue) or by MS1 only (red, GlycopeptideGraphMS with additional accurate-mass and isotopic pattern check of the raw data). The assignment of the compositions is based on information from all replicates. Lines between compositions indicate the mass difference for Hex (yellow), HexNAc (blue), HexHexNAc (green), Fuc (red), and NeuAc (purple). * Indicates potential deconvolution errors and ** indicates data not included in the Byonic search list.

Overall, the GlycopeptideGraphMS workflow showed a high identification accuracy (82/83, 99%) and coverage (82/83, 99%, **Table S4**, **Supporting Information File 1**). In comparison, the accuracy of the Byonic search for the glycopeptides of interest was comparably high (35/37, 95%), whereas the glycopeptide coverage was moderate (35/77, 45%). This is in line with the reported near-perfect accuracy and limited coverage of the glycopeptide identification by Byonic [85]. Of note, the glycopeptide coverage of Byonic depends highly on the search parameters, fragmentation settings, and the presence and quality of MS/MS spectra. The latter is often compromised due to dynamic range limitations, especially in complex matrices [85, 104]. The accuracy of both approaches (MS1 and MS/MS) may be impaired by unexpected peptide modifications, as exemplified for Cys oxidation. Thus, careful inspection of the result outputs (RT graphs in GlycopeptideGraphMS, automatically annotated MS/MS spectra in Byonic) is important. Indications of additional peptide modifications can then be considered for manual MS/MS verification and be included in the search space of automated MS/MS assignments in an iterative manner. Alternatively, a prior open search aimed at the identification of peptide modifications may be applied by software tools such as Preview [105]. Overall, this data shows that, while MS/MS-based assignment tools are essential for the confident identification of glycopeptide clusters, MS1-based approaches show a highly complementary performance by identifying glycopeptides for which no MS/MS data is present. For the latter, GlycopeptideGraphMS is a highly valuable tool as it is easy to use, fast, and open source.

2.2.3. GLYCOPEPTIDE CURATION AND QUANTIFICATION IN LACYTOOLS

Upon glycopeptide identification, the analytes were curated and quantified by LaCyTools. The performance of LaCyTools was compared to that of Skyline (manual curation and quantification) and GlycopeptideGraphMS (quantification). The analytes and charge states passing the quality criteria (for LaCyTools: *m/z* accuracy <10 ppm, isotopic pattern quality value <0.2, signal-to-noise ratio >9; for Skyline: *m/z* accuracy <10 ppm, idotp >0.85) were highly similar between LaCyTools and Skyline (**Table S5**, **Supporting Information File 1**). Minor differences were observed for low-abundant glycopeptides. In GlycopeptideGraphMS, quality control is only based on mass accuracy and not included in this comparison.

The three software tools evaluated for targeted glycoform quantification resulted in comparable site-specific glycosylation profiles for human plasma IgG, JC, and IgA2 (**Figure 2.4** and **Table S5**, **Supporting Information File 1**), which were in line with the literature (**Table S6**, **Supporting Information File 1**) [91, 92]. Skyline and LaCyTools showed the highest similarity in the relative quantification results (**Figure S20**, **Supporting Information File 2**). Both tools had a median relative standard deviation (RSD) of 4% over all quantified glycopeptides. In contrast, GlycopeptideGraphMS integration resulted in a higher variability (median RSD: 15%, **Figure S20**, **Supporting Information File 2**) and slightly deviating glycosylation profiles, as compared to Skyline and LacyTools.

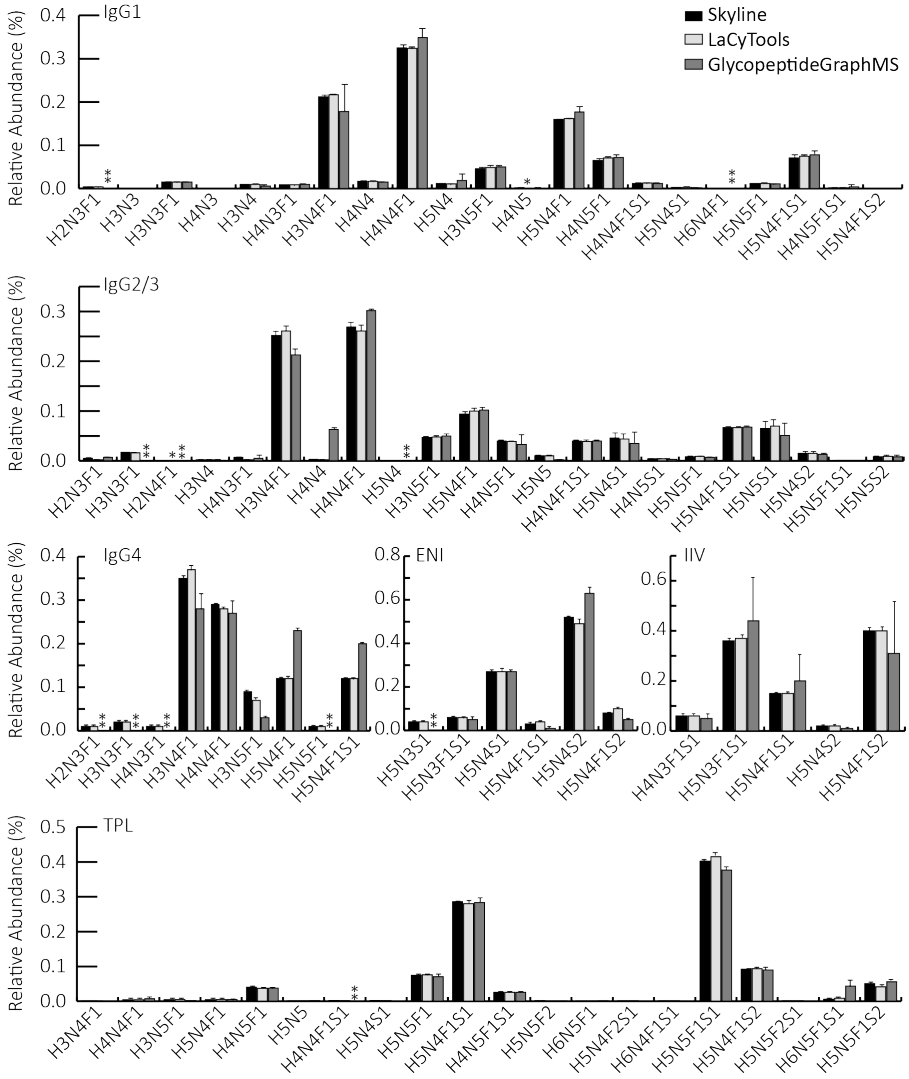

Figure 2.4: Comparison of quantification results obtained by manual integration of EICs in Skyline (black), automated integration of summed MS spectra in LaCyTools (light gray), and GlycopeptideGraphMS (dark gray). Error bars represent standard deviation of MS1-only measurements (n = 4 for LacyTools and Skyline; n = 3/4 for GlycopeptideGraphMS; in all detected replicates, n was at least 3. The first injection was excluded for all tools due to RT shifts and increased standard deviations). *: Did not pass the analyte curation (LaCyTools). **: Was not identified in at least 3 technical replicates (GlycopeptideGraphMS).

As the data used for quantification were the same, the differences in the quantification precision are caused by the data processing performed by the different software tools. Of note, the automated quantification in GlycopeptideGraphMS required additional manual interference for analytes that had multiple RTs in the output file and only a single chromatographic peak in the raw data. Similar as for the glycopeptide identification, quantification with GlycopeptideGraphMS showed clearly that the preprocessing of the data is a crucial factor for the outcome. Further optimization of the OpenMS preprocessing steps to prevent double feature assignments may improve the quantification precision.

Within the investigated quantification tools, Skyline allows the highest control of the feature selection for quantification as the integrated EICs can be manually inspected for interferences, correct peak integration, and quality criteria (mass accuracy and isotopic pattern). LaCyTools provides information on the mass accuracy and isotopic pattern and integrates the isotopes of selected features in summed MS spectra within user-defined RT windows. Here, it is crucial to select appropriate RT windows and isotopes of interest before starting the analysis to prevent the inclusion of closely eluting isomeric and isobaric interferences. Of note, isomeric glycopeptide compositions were summed and not processed individually. This approach makes RT alignment a crucial step for a robust quantification. With the optimized parameters in place, LaCyTools allows highly automated data handling, making it an excellent tool for, e.g., clinical cohort analysis. In the current work, a python script was developed to streamline the connection between GlycopeptideGraphMS identification and LaCyTools quantification (**Supporting Information File 3**). All tools provided absolute values for glycopeptide quantification, which were subsequently total-area-normalized per glycosylation site, as commonly done in label-free relative quantification in glycoproteomics [90, 91, 104, 106].

In the current study, all quantitative analyses were performed on the MS1-only runs to obtain the highest possible data quality. However, runs including fragmentation scans are also suitable for quantification, albeit introducing a slightly higher variability in some cases due to a lower number of data points per chromatographic feature (in particular obvious for the IgG1 and IgG2/3 data in the current study, see **Figure S21**, **Supporting Information File 2**). The difference in the quantification accuracy between MS1-only and MS/MS data is highly dependent on the frequency of the MS1 scans, and thus the time spent on fragmentation scans. In most situations, it is likely that a compromise must be made to allow both robust quantification and data-rich MS/MS identification in the same LC–MS run. The introduction of MS1-based identification reduces the time needed for fragmentation.

2.3. CONCLUSION

Here, we demonstrated a semiautomated glycoproteomics data analysis workflow for enriched glycoproteins by integrating different tools for glycopeptide identification, curation, and quantification after RPLC separation and MS(/MS) detection. For this, a mix of the human plasma-enriched antibodies IgG and IgA was used as a representative glycoproteomics sample of moderate complexity. A similar approach can be applied to a more complex sample when targeting only a select set of glycoproteins. However, to capture the full complexity of, e.g., the human glycoproteome, improvements should be made in the automated integration between the described tools. In line with previous reports on single glycoproteins, the number of identified glycoforms was significantly maximized by combining MS1-based identification (using GlycopeptideGraphMS) in combination with MS/MS-based identification (using Byonic) as compared to fragmen-tation-based analysis alone. Moreover, the graphical approach allowed by GlycopeptideGraphMS is very powerful for identifying unexpected glycoforms as well as modifications of the glycopeptides and aids the optimization of the search space for MS/MS annotation in an iterative manner. Although an MS1-based approach alone allows the identification of more unique glycopeptides as compared to an MS/MS-based approach, a combined workflow is essential to prevent wrongly assigned glycopeptides as well as to identify the nature of specific modifications. The combination of Byonic and GlycopeptideGraphMS identification with LaCyTools-based curation and quantification of glycopeptides from enriched glycoproteins as presented in the current work provides a powerful workflow towards high-throughput glycopeptide analysis.

2.4. EXPERIMENTAL

2.4.1. SAMPLE, CHEMICALS, AND ENZYMES

Human plasma Visucon-F was obtained from Affinity Biologicals (Ancaster, ON, Canada). Affinity matrix beads for IgG (CaptureSelect FcXL, capacity 25–35 g/L) and IgA (CaptureSelect IgA, capacity 8 g/L) were obtained from ThermoFisher Scientific (Leiden, Netherlands). All used chemicals were from Sigma-Aldrich (Zwijndrecht, Netherlands) except for trifluoroacetic acid (Merck, Darmstadt, Germany) and acetonitrile (Biosolve, Valkenswaard, Netherlands). Purified water was used from a Purelab Ultra system (Veolia Water Technologies Netherlands B.V., Ede, Netherlands). Sequencing-grade trypsin was obtained from Promega (Madison, WI).

2.4.2. SAMPLE PREPARATION

A detailed description of the methods for the immunoaffinity enrichment of the immunoglobulins and the glycopeptide preparation can be found elsewhere [91]. In brief, 5 µL of Visucon F plasma standard were diluted in PBS, and the immunoglobulins

were enriched using a mix of CaptureSelect FcXL Affinity matrix beads for IgG and CaptureSelect IgA affinity matrix beads for IgA. Upon incubating the serum and the beads for 1 h at room temperature with agitation, the beads were washed three times with PBS and three times with water. The immunoglobulins were released by acid elution (100 mM formic acid) and collected into a 96-well PCR plate (Greiner Bio-One, Kremsmünster, Austria). Finally, the eluates were dried for 2.5 h at 60°C by centrifugation under vacuum.

For tryptic digestion, the dried sample was reconstituted in 10 µL of reduction–alkylation buffer containing 100 mM Tris buffer, 1% w/v SDC, 10 mM tris(2-carboxyethyl)phosphine (TCEP), and 40 mM chloroacetamide (CAA). Upon mixing for 5 min, the samples were incubated for 5 min at 95°C and cooled to room temperature. Tryptic digestion was started by the addition of 50 µL digestion buffer containing 50 mM ammonium bicarbonate pH 8.5 and 200 ng sequencing-grade trypsin. Upon mixing for 5 min, the sample was incubated at 37°C overnight. Acid precipitation using 1.2 µL formic acid was performed on the following day. The precipitate was removed by centrifugation, and 40 µL of the supernatant was transferred to a V-bottom 96-well plate (Greiner). The sample was stored at -20°C.

2.4.3. LC–MS/MS ANALYSIS

A 0.5 µL aliquot of the sample was analyzed five times with MS1 only (for MS1-based identification in GlycopeptideGraphMS and quantification in LaCyTools, Skyline, and GlycopeptideGraphMS) and five times with additional MS/MS (for fragmentation-based identification using Byonic and quantification using LaCyTools) in an alternating order. For the separation of the (glyco)peptides, the sample was injected into an Easy nLC 1200 system (Thermo Fisher Scientific) equipped with an in-house prepared precolumn (15 mm × 100 µm; Reprosil-Pur C18-AQ 3 µm, Dr. Maisch, Ammerbuch, Germany) and an analytical nanoLC column (15 cm × 75 µm; Reprosil-Pur C18-AQ 3 µm). As mobile phases 0.1% formic acid in water (A) and 20% water/80% acetonitrile + 0.1% formic acid (B) were used. A gradient from 10–40% of the mobile phase B was applied within 20 min. The LC was hyphenated to an Orbitrap Fusion Lumos MS (Thermo Fisher Scientific). For MS1 analysis, scans were acquired in a mass range of *m/z* 400–3,500 in positive mode. The resolution was set to 120,000. The target for automatic gain control (AGC) was set to 400,000. The maximum injection time was 50 ms. An intensity threshold of 20,000 was applied. For MS/MS analysis, charge states 2–7 were included for stepped higher-energy C-trap dissociation (HCD) with a normalized collision energy (NCE) of 35% ± 5% (30%, 35%, and 40% combined in one spectrum), a maximum injection time of 60 ms, and a AGC target of 50,000. Additionally, MS/MS fragmentation was triggered for a HexNAc loss (204.087). For the triggered MS/MS analysis, a stepped HCD with an NCE of 35% ± 15% (20%, 35%, and 50% combined in one spectrum) was applied, and the AGC target was increased to 500,000 while the maximum injection time was increased to 200 ms. For all MS/MS scans, a precursor isolation width of *m/z* 1.2 was used. The MS/MS scan resolution was

30,000 and the *m/z* range was 110–3,500.

2.4.4. MS/MS DATA EVALUATION

A manual inspection of the raw data was performed in Xcalibur (v. 2.2, Thermo Fisher Scientific). PMI-Byonic (v. 3.7.13 Protein Metrics) was used for the MS/MS-based protein and glycosylation site identifications [88]. Protein identification was based on a canonical Homo sapiens UniProt database including 71,591 protein sequences (20,205 from Swiss-Prot and 51,386 from TrEMBL). The C-terminal cleavage of lysine and arginine and a maximum of two missed cleavages was allowed. A tolerance of 10 ppm was applied for the precursors and 20 ppm for fragment ions. A carbamidomethylation was set as a fixed modification for cysteine residues. Methionine oxidation was enabled as a variable modification. The search for *N*- and *O*-glycopeptides was separately performed. For this purpose, either the database "N-glycan 309 mammalian no sodium" (**Supporting Information File 7**) or "O-glycan 78 mammalian" (**Supporting Information File 6**) was applied as a custom modification. For manual MS/MS assignments, the web tool ProteinProspector v. 6.2.1 was used (`http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct`).
All glycopeptide compositions that were not identified by Byonic were subjected to a manual check of the MS/MS raw data in Xcalibur. This check included verifying the presence of the characteristic MS/MS ions (**Table S4**, **Supporting Information File 1**). In addition, allotypes of IgG3 and IgA2, which can be present in a human plasma pool [78], were manually checked. For this, the peptide sequences TKPWEEQYNSTFR, GFYPSDIAVEWESSGQPENNYNTTPPMLDSDGSFFLYSK (IgG3 *N*-glycopeptides), and MAGKPTHINVSVVMAEADGTC(Y) (IgA2 *N*-glycopeptide) were checked for the presence of the Y1 (peptide + HexNAc) ion in the MS/MS data. In addition, the expected glycoforms H1N1, H1N1S1, and H1N1S2 of the IgG3 *O*-glycopeptide SCDTPPPCPR were checked.

2.4.5. GLYCOPEPTIDEGRAPHMS ANALYSIS

MS1-based glycopeptide identification in all five MS1-only measurements and visualization was performed using GlycopeptideGraphMS (v. 2.06) according to the user manual [89]. In short, the raw data were first transformed to the mzML format using msconvert (ProteoWizard 3.0 suite). The data preprocessing included the deconvolution of all MS1 signals using an OpenMS workflow (KNIME_OPENMS_GraphMS_Preprocessing_120318) in KNIME [89, 107, 108]. This workflow was used with OpenMS 2.3. Adaptions in the parameters were made in the *m/z* range of 400–3,500 and the charge states 2–7. For the glycopeptide identification in GlycopeptideGraphMS, the intensity threshold was set to 1,000,000, the allowed mass deviation of the glycan building blocks to 0.02 Da, and the maximum subgroup degree was set to 1. As composition searching blocks (see the example provided in **Supporting Information File 5**), hexose (Hex, 162.0528 Da, max. 30 s RT difference).

N-Acetylhexosamine (HexNAc, 203.0794 Da, max. 30 s RT difference), hexose, and *N*-acetylhexosamine (HexHexNAc, 365.1322 Da, max. 30 s RT difference), deoxyhexose (Fuc, 146.0579 Da, max. 20 s RT difference), and *N*-acetylneuraminic acid (NeuAc, 291.0954 Da, max. 120 s RT difference) were enabled. For each glycopeptide cluster of interest, one data point was assigned to a composition that was verified by the Byonic search. For the visualization in GlycopeptideGraphMS, the diameter of the data points and the relative abundance of the glycopeptides were represented upon logarithmic scaling between intensities from 1×10^6 to 1×10^{12}. False-positive assignments containing negative values in the compositions (illogical compositions) based on the assigned reference data points of all glycopeptides were removed. Analytes (with logical compositions) connected solely to analytes with illogical compositions (i.e., negative features) were excluded as well. For quantitative comparisons, only analytes were considered which were identified in at least three technical replicates. Intensities of analytes present at more than one RT were summed in case of a close RT proximity (likely isomers) or manually checked in the raw data for multiple peaks and included or excluded, dependent on the presence of multiple peaks in the raw data.

2.4.6. SKYLINE ANALYSIS

In addition to the automated glycopeptide identification, a MS1 assignment and peak integration was performed in Skyline (v19.1.0.193). The correct peak integration was manually checked. A reference glycopeptide composition list was inserted into Skyline. This list contained the merged information from the automatically assigned compositions (Byonic and GlycopeptideGraphMS), compositions listed on GlyConnect [94] for IgG and IgA, and an in-house analyte list that was recently used for an IgG/IgA analysis (based on literature information and manual peak assignment in MS1) [91]. The transition settings were set to product ions, the charge states were set to 2–7, and the time window was adjusted for each different glycopeptide cluster. MS1 data of the glycopeptide compositions were manually inspected, and charge states with an isotope dot product (idotp) >0.85 and a mass accuracy <10 ppm were included. "Normalized Area" was used for quantification.

2.4.7. LACYTOOLS ANALYSIS

For automated quantification in LaCyTools (v 1.0.1) [90], the raw data were converted to the mzXML format by MSConvert. The generation of the LaCyTools analyte list was supported by an in-house Python (v 3.7.6) script (**Supporting Information File 3**), which converted a representative GlycopeptideGraphMS output to the required input format for LaCyTools. Glycopeptide compositions that were not assigned in the representative data set in GlycopeptideGraphMS were added to the list to an appropriate retention time cluster. Potentially false-positive results (no MS1 isotope pattern matching or no MS/MS verification) were manually removed. The applied analyte list is provided in **Supporting Information File 5**. Next, an alignment list was

created by selecting the most abundant glycopeptide compositions for each RT cluster. The width of the retention time cluster was set to 15 s and adjusted to 7 s for analytes with closely eluting interference signals. The RT alignment of the technical replicates was performed within a time window of 30 s and an *m/z* window of 0.1. For analyte curation and quantification, an *m/z* window of 0.025 was used. Upon processing in LaCyTools, all charge states of analytes with an isotopic pattern quality value higher than 0.2, mass accuracies of >10 ppm, and a signal-to-noise ratio <9 were excluded. The peak areas of the remaining charge states were summed and corrected by being divided by the isotopic pattern fraction. Of note, for the comparison of the relative quantification of GlycopeptideGraphMS, Skyline, and LaCyTools, the relative abundance was not renormalized to the intersection of the analytes.

SUPPORTING INFORMATION

Supporting information is available free of charge via `https://www.beilstein-journals.org/bjoc/articles/16/253`.

ACKNOWLEDGEMENTS

The authors would like to thank Paul J. Hensbergen for critically reading the manuscript. This research was funded by the European Commission H2020 (Analytics for Biologics project, Grant No. 765502).

3

Cysteine Aminoethylation Enables the Site-Specific Glycosylation Analysis of Recombinant Human Erythropoietin using Trypsin

Steffen Lippold,[1] Alexander Büttner,[2] Matthew S.F. Choo,[3] Michaela Hook,[2] Coen J. de Jong,[1] Terry Nguyen-Khuong,[3] Markus Haberger,[2] Dietmar Reusch,[2] Manfred Wuhrer[1] and Noortje de Haan[1]

[1]Center for Proteomics and Metabolomics, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, Netherlands
[2]Pharma Technical Development, Roche Diagnostics GmbH, 82377 Penzberg, Germany
[3]Bioprocessing Technology Institute, Agency for Science Technology and Research, Singapore 138668

Reprinted and adapted from Anal. Chem. 2020, 92, 14, 9476–9481
DOI: 10.1021/acs.analchem.0c01794

Recombinant human erythropoietin (rhEPO) is an important biopharmaceutical for which glycosylation is a critical quality attribute. Therefore, robust analytical methods are needed for the in-depth characterization of rhEPO glycosylation. Currently, the protease GluC is widely established for the site-specific glycosylation analysis of rhEPO. However, this enzyme shows disadvantages, such as its specificity and the characteristics of the resulting (glyco)peptides. The use of trypsin, the gold standard protease in proteomics, as the sole protease for rhEPO is compromised, as no natural tryptic cleavage site is located between the glycosylation sites Asn24 and Asn38. Here, cysteine aminoethylation using 2-bromoethylamine was applied as an alternative alkylation strategy to introduce artificial tryptic cleavage sites at Cys29 and Cys33 in rhEPO. The (glyco)peptides resulting from a subsequent digestion using trypsin were analyzed by reverse-phase liquid chromatography-mass spectrometry. The new trypsin-based workflow was easily implemented by adapting the alkylation step in a conventional workflow and was directly compared to an established approach using GluC. The new method shows an improved specificity, a significantly reduced chromatogram complexity, allows for shorter analysis times, and simplifies data evaluation. Furthermore, the method allows for the monitoring of additional attributes, such as oxidation and deamidation at specific sites in parallel to the site-specific glycosylation analysis of rhEPO.

3.1. INTRODUCTION

RECOMBINANT human erythropoietin (rhEPO) is a successful therapeutic glycoprotein whose bioactivity and safety are highly affected by specific glycosylation features (e.g., the rhEPO half-life depends on the number of sialic acids) [11]. First-generation rhEPO exhibits three *N*-glycosylation sites (Asn24, Asn38, Asn83) and one *O*-glycosylation site (Ser126). *N*-Glycans of rhEPO show a wide structural diversity, including a varying number of antennae, *N*-acetyllactosamine (LacNAc) repeats, terminating sialic acids (*N*-acetylneuraminic acid (Neu5Ac) and *N*-glycolylneuraminic acid (Neu5Gc)), and glycan modifications, such as acetylation, phosphorylation, and sulfation [77]. The *O*-glycosylation site reportedly carries core 1-type structures with 0, 1, or 2 sialic acids, including acetylation [109]. Numerous mass spectrometry (MS)-based analytical methods have been described to characterize rhEPO glycosylation at the released glycan level [77, 110], by studying glycopeptides [59, 111–114], or, more recently, analyzing rhEPO in its intact form [59]. Between those methods, the bottom-up glycopeptide analysis is the only approach that allows the simultaneous characterization of *N*- and *O*-glycosylation heterogeneity in a site-specific manner [11].

Site-specific glycopeptide analysis requires the generation of proteolytic peptide moieties covering a single glycosylation site [115]. Therefore, protease selection is a crucial step in analytical method development. Trypsin is the gold standard in proteomics due to its high specificity and robustness [116]. However, the use of this protease alone is not sufficient for site-specific glycosylation analysis specifically for rhEPO, as no natural tryptic cleavage sites (Lys and Arg) are located between the glycosylation sites Asn24 and Asn38 [11, 111]. GluC is currently the protease of choice in most rhEPO glycopeptide studies [111–114]. However, the use of this enzyme is generally compromised by its pH dependent specificity, its low digestion efficiency, and the generation of relatively large peptide portions [68]. Hence, numerous missed cleaved products and inconsistent results for the obtained peptide moieties are reported, hampering the data analysis of rhEPO glycopeptides [59, 111, 113, 114]. To overcome the disadvantages of the generation of rhEPO glycopeptides by GluC, different approaches were established. For example, trypsin and GluC were used in a double digestion of rhEPO[117] or information from two independent complementary digestions was combined [59, 114]. However, these strategies are less consistent and more time-consuming than the use of a single specific protease.

In the current study, site-specific glycosylation analysis of rhEPO using trypsin as the sole proteolytic enzyme was enabled via the aminoethylation of cysteines. Using this approach, cysteines are transformed into pseudolysines and recognized as substrates by trypsin [118]. Aminoethylation can be performed using 2-bromoethylamine, as was first described in 1956 [119]. This approach was applied by several others to facilitate proteolytic cleavage using trypsin [120, 121]. For example, cysteine aminoethylation and trypsin digestion were recently combined in a biopharmaceutical application, increasing the sequence coverage of the complementarity determining region in a

monoclonal antibody [121]. In the current study on rhEPO, cysteine aminoethylation created two tryptic cleavage sites between Asn24 and Asn38 at Cys29 and Cys33. This approach was readily integrated into a conventional reverse-phase liquid chromatography (RP-LC)-MS workflow for rhEPO multiple-attribute monitoring at the (glyco)peptide level. As compared to conventional methods, the current approach resulted in a lower diversity in the cleavage products, shorter peptide portions for the glycopeptides, and shorter analysis times. Furthermore, it prevented the interference of oxidation (Met54) and deamidation (Asn47) sites with the assessment of the Asn38 glycosylation.

3.2. Experimental Section

3.2.1. Chemicals and Samples

All chemicals had at least analytical grade quality. Further information about the chemicals used can be found in the **Information S1**. A reference standard of rhEPO produced in a Chinese hamster ovary (CHO) cell line was provided by Roche Diagnostics (Penzberg, Germany).

3.2.2. Sample Preparation

A detailed description of the sample preparation can be found in the **Information S1**. Briefly, 250 µg of rhEPO was denatured, reduced (20 mM DTT), and then alkylated by either 60 mM iodoacetic acid (for GluC digestion as reported [122]) or 60 mM 2-bromoethylamine (for 1 h at 60 °C for trypsin digestion) in a final volume of 320 µL. Upon buffer exchange, samples were digested for 16-18 h by GluC (25 °C, enzyme/protein 1:25) or trypsin (37 °C, enzyme/protein 1:100) in 450 µL of 50 mM ammonium bicarbonate, pH 7.8. Sample preparation was performed in triplicate per digestion protocol.

3.2.3. Reverse-Phase Liquid Chromatography-Mass Spectrometry

Ten µL of each sample (around 5 µg) was analyzed by C8 RP-LC-MS/MS. A C8 BEH column (2.1 mm × 150 mm, 1.7 µm, 130 Å, Waters) was used on a UPLC system (Vanquish Horizon, Thermo Scientific). The flow rate was kept at 300 µL/min, and the column temperature was kept at 65 °C. A 65 min gradient of 1% to 80% B (mobile phase A, 0.1% formic acid in water; mobile phase B, 0.1% formic acid in acetonitrile) was applied (**Table S1**). The UPLC was hyphenated to an LTQ Orbitrap Velos (Thermo Scientific). MS1 information was obtained for *m/z* 200-2,000 in the positive ion mode. MS2 scans were acquired by CID with a normalized collision energy of 35% for the five most intense parent ions. Detailed information about the MS settings can be found in **Information S2** [122].

3.2.4. Data Analysis

A detailed overview of the data analysis workflow can be found in **Information S3** and **Figure S1**. In short, GlycopeptideGraphMS (ver. 2.04) [89], in combination with the MS2 confirmation (manual and using Byonic) of selected glycoforms (**Figure S2**) and the literature [122], was used for assignment of the glycopeptides. An initial analyte quality control and automatic quantification of the main cleavage products was performed in LacyTools (ver. 1.0) [90]. Skyline (ver. 19.1.0.193) was used for a manual integration and the quality control of the main glycopeptides of interest for all relevant cleavage products [123].

3.3. Results and Discussion

3.3.1. Proteolytic Peptides

All expected peptides, with at least three amino acids, were assigned for the tryptic rhEPO digest (**Figure S3** and **Table S2**). The efficiencies of the Cys alkylation with 2-bromoethylamine (97.3%) and the subsequent tryptic digestion (99.3%) were evaluated based on one Cys-containing peptide with an adjacent Arg (**Figure S4** and **Table S3**). The alkylation selectivity was evaluated by considering the top seven most abundant peptides, and showed negligible rates (<1%) of Met, Glu, Asp, and Tyr aminoethylation (**Figures S5-S10** and **Table S4**). The non-selectively alkylated sites are in line with the reported (*O*-)alkylation byproducts in *S*-alkylation reactions [124]. Trace amounts of peptides were observed of which the presence is likely explained by the tryptic cleavage of alkylated Met, Glu, or Tyr (**Figures S7** and **S9**). For the glycopeptides obtained in the trypsin workflow, cleavages were observed C-terminally to Lys (Asn38 and Asn83), Arg (Ser126), and aminoethylated Cys (Asn24) (**Figure 3.1** and **Table S5**). Aminoethylation of the cysteines and their cleavage by trypsin was confirmed by fragmentation of the non-glycosylated Asn24 (theoretical *m/z* 980.4359, 1+) and Asn38 (theoretical *m/z* 665.8466, 2+) peptides (**Figure S11**). For both Asn24 and Asn38, an additional missed cleaved peptide portion was detected, which was overalkylated. The location of the extra aminoethyl group was assigned to Cys29 in the MS2 analysis of the non-glycosylated Asn24 peptide Glu21–Cys31 (**Figure S12**), likely preventing trypsin cleavage. The estimated relative abundance of the overalkylated and missed cleaved versions of the Asn24 and Asn38 glycopeptides was 9.1% and 10.5%, respectively (**Table S6**). Considering the theoretically increased ionization efficiency of the missed cleaved version, due to the additional amines in the elongated sequence [125], the true relative abundance may be even lower and consequently negligible. For the other glycosylation sites (Asn83 and Ser126), specific peptide portions were obtained using trypsin (**Figure 3.1**).

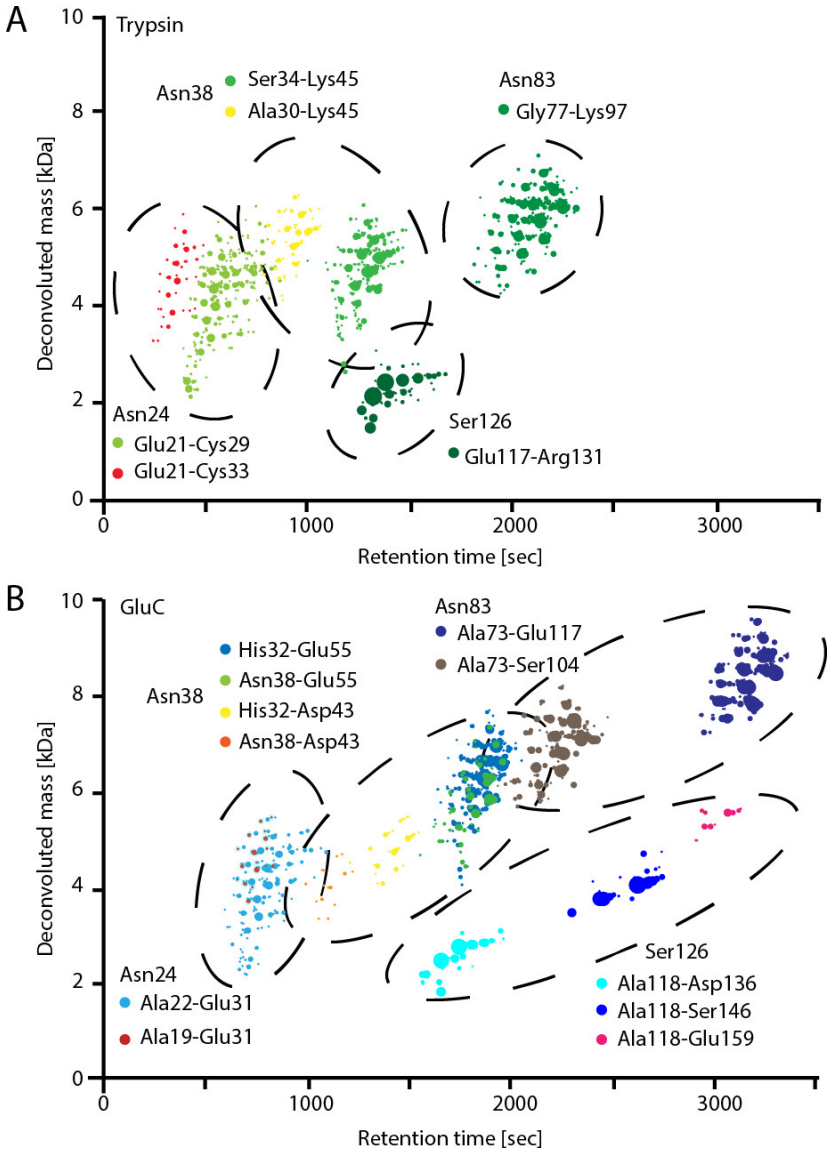

Figure 3.1: Visualization of detected rhEPO glycopeptides in representative replicates by RP-LC-MS/MS analysis upon (**A**) trypsin digestion after aminoethylation and (**B**) GluC digestion after carboxymethylation. Glycopeptide assignment and visualization was performed using GlycopeptideGraphMS [89]. The diameter of the data points indicate the relative abundance of the glycopeptides (logarithmic scaling between intensities from 1×10^6 to 1×10^{12}). Data points with the same color have the same peptide backbone but different glycan compositions.

GluC-derived rhEPO glycopeptides showed C-terminal cleavages to Glu (Asn24, Asn38, Asn83, and Ser126), Asp (Asn38 and Ser126), and Ser (Asn83 and Ser126), which were reported earlier for rhEPO [111]. The unspecific cleavage of GluC resulted in several digestion products per glycosylation site. Missed cleaved products of Asn24 (Ala19–Glu37) and Asn38 (Asn38–Glu55) resulted in coeluting glycopeptides in the same cluster as the respective main peptide moieties (**Figure 3.1**), hampering the data analysis [59, 114]. The highest heterogeneity in the obtained GluC peptide portions was observed for Asn38, which showed four cleavage products (**Figure 3.1** and **Table S5**), as reported previously [114]. Furthermore, Ser126 was detected on three different peptide portions, while the GluC digestion resulted in two products for both Asn24 and Asn83. The abundance of additional peptides ranged from 0.3% for Asn38 (Asn38–Asp43) to 29.2% for Ser126 (Ala118–Ser104) (**Table S6**). For further data analysis, peptides of interest were selected based on the relative abundance (>20%). In addition, a GluC cleavage product of Asn38 (Asn38–Glu55) was included, as its glycosylation site is located next to the cleavage site. The rhEPO cleavage products of GluC are not consistently reported in different studies [59, 111, 113, 114]. This may be due to different protease sources, sample preparations, and data analysis procedures.

Using the novel trypsin workflow, all glycopeptide peptide moieties, including the missed cleaved products for Asn24 and Asn38, were well separated by RP-LC, which allowed for straightforward data integration. Furthermore, the trypsin workflow enabled a shorter analysis time as compared to the GluC workflow using the same RP-LC gradient. The maximum retention time for the trypsin workflow was 37 min (Asn83-H7N6F1S4), while that of GluC was 55 min (Asn83-H7N6F1S4). In summary, the trypsin workflow reduced the heterogeneity of proteolytic glycopeptides as compared to the GluC workflow (6 vs 11 glycopeptides for four glycosylation sites) because of the higher protease specificity of trypsin.

3.3.2. Site-Specific Glycosylation Analysis

In line with previous reports on rhEPO glycosylation, the trypsin-based workflow showed mainly fucosylated complex-type *N*-glycans with varying numbers of antennae (2-4), LacNAc repeats (1-3), and sialylation (0-5) (**Table S7**) [77, 109, 110]. Moreover, site-specific glycosylation differences were characterized, indicating a relatively high content of high mannose-type and diantennary complex-type structures on Asn24; Asn38 and Asn83 predominantly showed tetra-antennary complex-type *N*-glycans (**Figure 3.2**) [77]. For the *O*-glycosylation site, mainly core 1-type structures with one and two sialic acids were detected, which was in line with those reported in the literature [109]. Low abundances of Neu5Gc-containing glycoforms were assigned for both *N*- and *O*-glycopeptides (**Table S7**). Interestingly, glycopeptide compositional isomers were separated in the RP-LC (**Figure S13**). However, the isomers were not individually considered for further data processing due to uncertainties in their assignment. A detailed exploration of the isomeric structure assignment was considered out of the scope of this report.

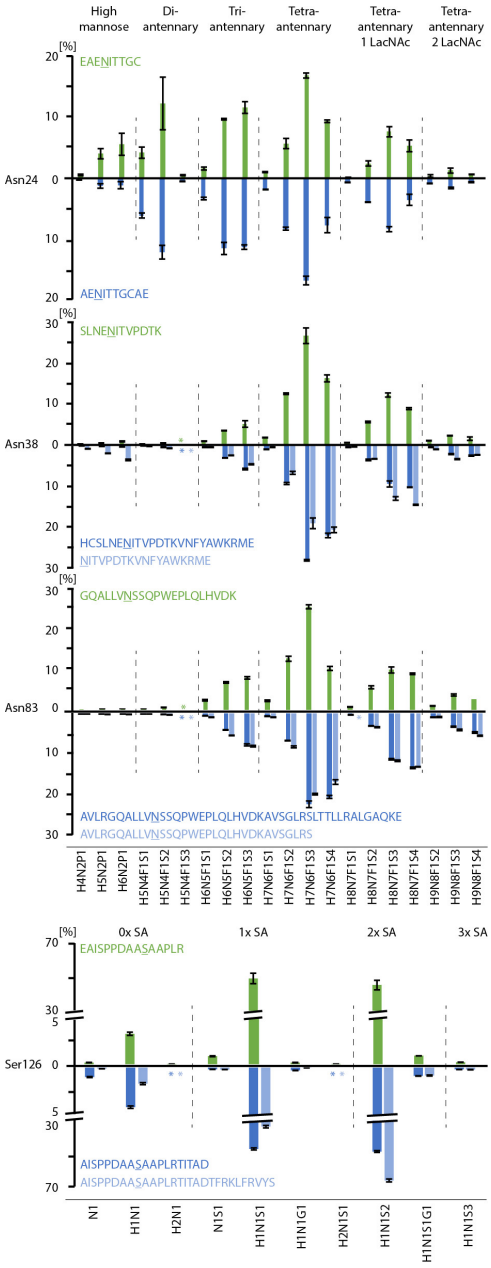

Figure 3.2: Site-specific quantitative comparison of the main *N*- and *O*-glycoforms (>0.5% relative abundance excluding acetylated variants, **Table S7**) of the peptide moieties of interest obtained by trypsin (green) or GluC (blue) digestion of rhEPO. Bars and error bars represent the mean values and standard deviations of triplicate measurements, respectively. Asterisks (*) indicate nondetected glycoforms. Compounds are represented as H, hexose; N, *N*-acetylhexosamine; F, fucose; S, *N*-acetylneuraminic acid; G, *N*-glycolylneuraminic acid; and P, phosphate. SA: sialic acid (S or G).

A total number of 39 (74), 32 (56), 28 (58), and 10 (20) compositions for the main trypsin-derived Asn24, Asn38, Asn83, and Ser126 glycopeptide clusters, respectively, were assigned, excluding (including) acetylation variants (**Table S7**). This amounts to 55 (100 including acetylation) different *N*-glycoform compositions. Both *N*- and *O*-glycopeptides revealed low abundant glycoforms, which were previously not reported for rhEPO. For example, hybrid-type structures were detected on Asn24, and a disialyl motif (H1N1S3) was detected on Ser126 (for MS2, see **Table S7** and **Figure S14**). Of note, the glycopeptide feature assignment was based on the retention time clustering using GlycopeptideGraphMS and prior knowledge [89, 122]. To limit the assignments of false-positive glycoforms, MS2 validation was performed for at least one glycoform per cluster (**Figure S2**), and all MS1 signals were subjected to accurate mass and isotopic pattern matching. Glycoforms that have not been reported in CHO cell-produced rhEPO were in part validated by MS2 (**Figure S14**). The new trypsin method resulted in similar site-specific glycosylation profiles as compared to the reference GluC method (**Figure 3.2**). Interestingly, the relative abundance of high mannose glycans at Asn38 was higher for GluC peptide moieties, which had the cleavage site adjacent to the glycosylation site (Asn38–Glu55). Furthermore, differences in the relative abundances of GluC *O*-glycopeptides carrying 46% or 33% H1N1S1 and 47% or 64% H1N1S2 were observed on the shorter or longer peptide, respectively. The findings for Asn38 suggest a glycoform-dependent preference of the GluC digestion when the glycan is located subsequent to the cleavage site. Such effects were shown before for trypsin, which cleaved less efficiently when a fucosylated glycan was present next to the targeted Arg [96]. For the *O*-glycopeptides, the effects might be attributed to differences in MS responses between different cleavage products. To our knowledge, no reports have as of yet investigated the differences in the relative quantification of glycosylation for differently cleaved peptide moieties in rhEPO. Here we show that tryptic peptides result in a simplified data interpretation as compared to GluC, as less peptide heterogeneity is introduced.

In addition to the different glycoforms, the non-glycosylated variants of each of the tryptic cleavage products of Asn24 (2.3%), Asn38 (0.001%), and Ser126 (5.1%) were also detected (**Table S8**). It should be noted that the estimation of the abundance of the non-glycosylated variants would be more accurate after deglycosylation, as the ionization efficiency of the peptide with and without glycan may be vastly different, as previously reported [13]. The GluC-derived peptides showed inconsistent results for non-glycosylated variants on different cleavage products, in particular when the cleavage site was adjacent to the glycosylation site (Asn24 and Asn38). For example, non-glycosylated variants were only detected for Asn38 peptide moieties with the glycosylation site adjacent to the cleavage site (**Table S8**). In addition, the non-glycosylated variant of NITTGCAE was observed for the Asn24 site in the GluC digest (**Figure S15**). However, no glycopeptides were observed for this peptide moiety. Overall, the trypsin workflow facilitated a consistent site occupancy determination.

3.3.3. Oxidation and Deamidation

In addition to glycosylation, the post-translational oxidation (Met54) and deamidation (Asn47 and Asn147) of rhEPO were previously reported to impact its structure and biological activity [126]. Trypsin-based RP-LC-MS peptide mapping to assess the oxidation and deamidation of rhEPO is preferred over GluC-based methods, as trypsin allows the analysis of Asn47 deamidation and Met54 oxidation on separate peptides (**Figure 3.3**, **S16-S18**, and **Table S9**) [126]. GluC cleavage faces the issue of Met54 and Asn47 being present on the same peptide moiety as Asn38. This complicates the quantification of the individual post-translational modifications, as both the oxidation and the deamidation result in additional peaks per glycoform. Additionally, it results in ambiguities as the mass difference between a Neu5Gc and a Neu5Ac is exactly the same as the mass increment upon Met54 oxidation (15.9949 Da). Thus, besides site-specific glycosylation analysis, the new trypsin-based workflow offers an approach for multiple-attribute monitoring of rhEPO, including its deamidation and oxidation.

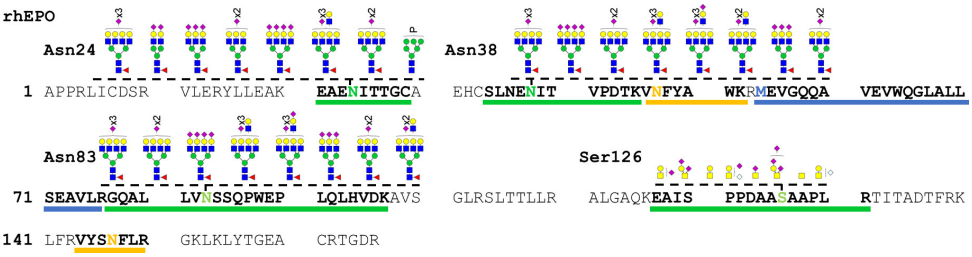

Figure 3.3: Sequence of rhEPO. Tryptic peptides covering glycosylation sites (green) and relevant deamidation (orange) and oxidation sites (blue) are highlighted. For the glycosylation sites, the eight most abundant glycoforms are presented. Structural ambiguities of the glycoforms were reduced by information from the literature [77].

3.4. Conclusions

We demonstrated the feasibility and advantages of cysteine aminoethylation with subsequent trypsin digestion for the site-specific glycosylation analysis of rhEPO. To our knowledge, this is the first report that describes trypsin as the sole protease for the cleavage between the glycosylation sites Asn24 and Asn38 of rhEPO. Trypsin showed a high specificity, resulting in six peptide portions covering the four glycosylation sites. In contrast, GluC showed a total of 11 peptide portions for the four glycosylation sites. In the direct comparison between carboxymethylated GluC-derived and aminoethylated trypsin-derived glycopeptides, we found comparable site-specific relative quantification results for the glycosylation of rhEPO. It should be noted that the current study is of exploratory nature, and further validation is needed to assess the quantification consistency of low-abundant glycoforms. The decreased heterogeneity in the proteolytic cleavage products using trypsin resulted in a significantly reduced

sample complexity, causing less ambiguities and facilitating straightforward data analysis. Moreover, the tryptic digestion did not show glycosylation-dependent cleavage and appeared highly suitable for the parallel monitoring of deamidation and oxidation in a multiple-attribute monitoring approach. Finally, the described sample preparation can be integrated in current rhEPO glycopeptide mapping workflows with minor adaptations, making it an attractive method for the biopharmaceutical sector.

SUPPORTING INFORMATION

Supporting information is available free of charge via `https://pubs.acs.org/doi/10.1021/acs.analchem.0c01794?goto=supporting-info`.

ACKNOWLEDGEMENTS

The authors would like to thank Peter van Veelen for his input on cysteine aminoethylation. Peter van Veelen and Arnoud de Ru are acknowleged for their support with the Byonic searches. This research was funded by the European Commission H2020 (Analytics for Biologics project, Grant 765502).

4

GLYCOFORM ANALYSIS OF INTACT ERYTHROPOIETIN BY MALDI FT-ICR MASS SPECTROMETRY

Steffen Lippold,[1] Raashina Thavarajah,[1] Dietmar Reusch,[2] Manfred Wuhrer[1] and Simone Nicolardi[1]

[1]Center for Proteomics and Metabolomics, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, Netherlands
[2]Pharma Technical Development, Roche Diagnostics GmbH, 82377 Penzberg, Germany

Reprinted and adapted from Anal Chim Acta 2021, 339084
DOI: 10.1016/j.aca.2021.339084

Recombinant human erythropoietin (EPO) is a complex therapeutic glycoprotein with three N- and one O-glycosylation sites. Glycosylation of EPO influences its safety and efficacy and is defined as a critical quality attribute. Thus, analytical methods for profiling EPO glycosylation are highly demanded. Owing to the complexity of the intact protein, information about EPO glycosylation is commonly derived from released glycan and glycopeptide analysis using mass spectrometry (MS). Alternatively, comprehensive insights into the glycoform heterogeneity of intact EPO are obtained using ESI MS-based methods with or without upfront separation of EPO glycoforms. MALDI MS, typically performed with TOF mass analyzers, has been also used for the analysis of intact EPO but, due to the poor glycoform resolution, has only provided limited glycoform information. Here, we present a MALDI FT-ICR MS method for the glycosylation profiling of intact EPO with improved glycoform resolution and without loss of sialic acid residues commonly observed in MALDI analysis. Three EPO variants were characterized in-depth and up to 199 glycoform compositions were assigned from the evaluation of doubly-charged ions, without any deconvolution of the mass spectra. Key glycosylation features such as sialylation, acetylation, and N-acetyllactosamine repeats were determined and found to agree with previously reported data obtained from orthogonal analyses. The developed method allowed for a fast and straightforward data acquisition and evaluation and can be potentially used for the high-throughput comparison of EPO samples throughout its manufacturing process.

4.1. INTRODUCTION

RECOMBINANT human erythropoietin (EPO) is a complex glycoprotein and a very successful biopharmaceutical [127–129]. The protein backbone is about 18,000 Da with additional glycosylation that extends the molecular mass up to about 32,000 Da. Four glycosylation sites (Asn24, Asn38, Asn83, and Ser126) and a high diversity of structural glycosylation features (varying levels of antennae, *N*-acetyllactosamine (LacNAc) repeats, sialylation, phosphorylation, sulfation, acetylation) have been described [130]. Since these features affect the safety and efficacy profile of EPO, glycosylation – defined as a critical quality attribute (CQA) of EPO – must be characterized to ensure the quality of the biopharmaceutical [122].

The in-depth glycosylation analysis of complex glycoproteins, such as EPO, is typically performed by a combination of different mass spectrometry (MS)-based methods [131, 132]. In fact, since the very first reports on the structure and function of EPO, MS has been widely used to study its glycosylation. The integration of information, obtained from multiple analysis levels, is needed for a comprehensive characterization of EPO glycosylation (**Figure 4.1**).

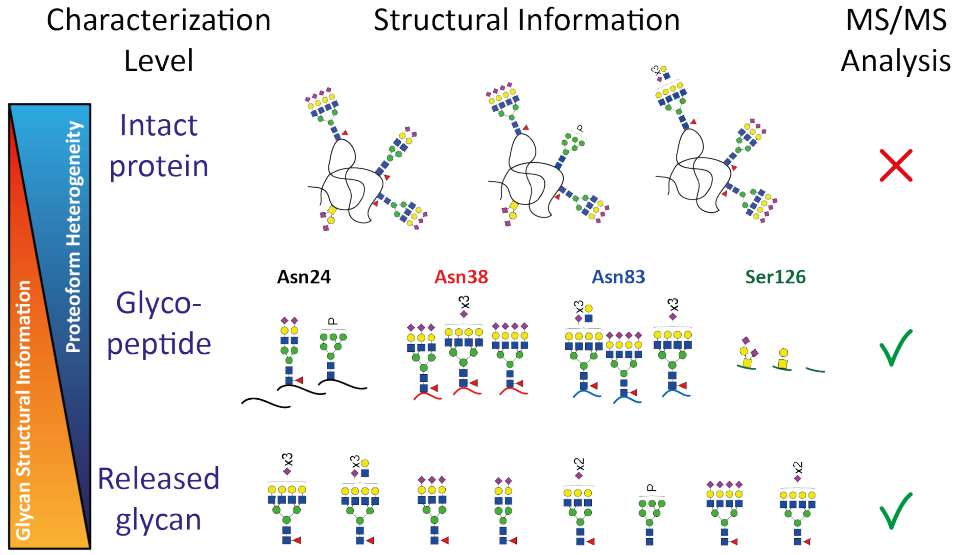

Figure 4.1: Scheme of multi-level characterization of EPO using MS-based methods.

The analysis of enzymatically released glycans has provided detailed information of the *N*-glycosylation of EPO [130, 133–136], while the analysis of glycopeptides has allowed for a site-specific determination of the glycan compositions [114, 115, 122, 137–140]. At the intact protein level, MS provides unique information on the combinatory presence of different glycan moieties on its four glycosylation sites [59, 133, 141–144]. However,

such an analysis, that leads to very complex mass spectra, is challenged by the high glycoform heterogeneity. This problem has been often mitigated by combining powerful separation techniques with sensitive MS methods [142, 145–148]. For example, the use of anion exchange chromatography (AEX) and Fourier transform ion cyclotron resonance (FT-ICR) MS has recently allowed for the assignment of 357 proteoforms to an EPO reference standard (EPO RS) [71]. Such an in-depth analysis requires long measurement times and tedious data processing, therefore, more direct strategies have also been developed. For example, in a recent report by Čaval and co-workers, intact glycoengineered variants of EPO were analyzed by native ESI MS on a modified Orbitrap MS system without any upfront separation [46]. The assignment of up to 236 EPO glycoforms allowed for comprehensive glycoform profiling of EPO [59]. Direct analysis methods were demonstrated to be very powerful for fast glycoform profiling and similarity assessment. Of note, the high heterogeneity of EPO limits the applicability of top-down MS/MS strategies therefore the analysis of intact EPO proteoforms, as well as other glycoproteins of similar complexity, must be always complemented by other structural analyses such as bottom-up MS.

Intact protein analysis of EPO has also been performed by MALDI MS however, this type of analysis has only provided limited information since the different EPO glycoforms are typically not resolved in time-of-flight (TOF) MS that is the most commonly used type of MALDI MS [149–152]. Furthermore, MALDI analysis of sialylated glycans and glycoconjugates is often affected by the loss of sialic acid residues that occurs during or post ionization. Compared to ESI MS, MALDI MS methods offer the advantage of a faster acquisition, the absence of carryover, a more stable ion source performance, and a more straightforward data interpretation. Therefore, a MALDI MS method for the analysis of intact EPO, with glycoform resolution and without loss of sialic acid residues, would be of great benefit.

In this study, we developed a MALDI FT-ICR MS method for the analysis of intact EPO glycoforms. Three EPO variants, namely EPO RS, EPO RP+ and EPO RP- were analyzed using 2,5-dihydroxyacetophenone as a MALDI matrix to minimize the loss of sialic acid residues. Robust and detailed glycosylation profiles were obtained directly from the evaluation of doubly charged EPO glycoforms, without any deconvolution of the mass spectra. The precise assignment of glycoform compositions allowed for the evaluation of key glycosylation features of EPO such as sialylation, acetylation, and LacNAc repeats. The developed method allowed for a straightforward comparison of EPO variants and can be potentially used throughout the manufacturing process for example for clonal selection, batch-to-batch comparison, and biosimilars evaluation.

4.2. Experimental Section

4.2.1. Chemicals and Samples

All chemicals used in this study had at least analytical grade quality and were, if not otherwise stated, purchased from Sigma Aldrich (Steinheim, Germany). EPO RS, EPO RP+, and EPO RP- were produced in Chinese hamster ovary (CHO) cells and were kindly provided by Roche Diagnostics (Penzberg, Germany). EPO RP+ and EPO RP- were obtained from EPO after fractionation by reversed-phase HPLC, as described elsewhere [122]. All EPO materials were previously studied by several methods and reference data was used for comparison [71, 122, 140]. Precisely, the very same samples used in this studies were previously characterized by bottom-up MS at a glycopeptide level and the obtained MS/MS spectra are available at doi:10.25345/C54998 [140]. For MALDI MS analysis, 2,5-dihydroxyacetophenone (DHAP, saturated) mixed in a ratio 8:2 with diammonium hydrogen citrate (DAHC, 20 mg/mL), in acetonitrile (ACN):water (50:50%, v/v), was used as MALDI matrix. 4-Chloro-α-cyanocinnamic acid (ClCCA) was prepared at 10 mg/mL in ACN:water (70:30%, v/v) while, saturated sinapic acid (SA) and saturated α-cyano-4-hydroxycinnamic acid (CHCA) were prepared in ACN:water:trifluoroacetic acid (30:69.9:0.1%, v/v/v). For external calibration of the MS system, 10 mM trans-2-[3-(4-tert-butylphenyl)-2-methyl-2-popenylidene]malononitrile (DCTB) in acetone was used as MALDI matrix to analyze cesium iodide cluster ions $[(CsI)_n + Cs]^+$.

4.2.2. Desialylation

Enzymatic desialylation of EPO RS was performed using SialEXO® (Genovis). 50 µg of EPO were digested with 50 units SialEXO (40 units/µL) in 20 mM tris(hydroxymethyl)aminomethane buffer, pH 6.8 at 37°C for 4 h.

4.2.3. Sample clean-up

The EPO samples were subjected to solid-phase extraction (SPE) by C4 ZipTip® tips (Merck Millipore). The SPE-tip was washed three times with 10 µL of ACN:water:formic acid (FA) (50:49.9:0.1%, v/v/v). Then, the SPE-tip was conditioned, three times, with 10 µL of 0.1% FA in water. Prior to loading, 10 µg of the sample were diluted in 10 µL with 0.1% FA in water. Upon sample loading (20 times), an additional washing step (five times methanol:water:FA (5:94.9:0.1%, v/v/v)) was applied and the sample was eluted in 10 µL of ACN:water:FA (50:49.9:0.1%, v/v/v).

4.2.4. MALDI spotting

0.5 µL of the EPO samples were spotted onto a polished steel target MALDI plate (Bruker Daltonics) and 0.5 µL of DHAP matrix were added. For external calibration of the MS system, 0.5 µL CsI_3 (50 mM in water) was spotted together with 0.5 µL of DCTB matrix [153]. The spots were allowed to dry at room temperature.

4.2.5. MALDI FT-ICR MS

A 15 T solariX XR FT-ICR mass spectrometer (Bruker Daltonics) with a CombiSource and a ParaCell was used in the study. The mass spectrometer was operated using ftmsControl software (Bruker Daltonics). MALDI experiments were performed with a Smartbeam-II Laser System (Bruker Daltonics) at a frequency of 500 Hz and 200 laser shots. The "selective accumulation" mode was used to sum 200 high-quality spectra. The data were acquired in positive-ion mode in the *m/z* range 2,022- 35,000 with 64,000 data points (i.e. transient time 0.2884 s). The total measurement time was approximately 5 min. The MS instrument was externally calibrated using $[(CsI)_n + Cs]^+$ ion clusters (**Figure S1**). All samples were measured in triplicates. For EPO RS, two additional measurements were performed on two days for interday analysis. Additionally, linear positive-mode MALDI TOF MS measurements were performed on a rapifleX mass spectrometer (Bruker Daltonics).

4.2.6. Data Analysis

Mass spectra were visualized and inspected in DataAnalysis 5.0 (Bruker Daltonics). mMass was also used for visualization and peak integration of .xy files [154]. The Uniprot sequence of human EPO (P01588) without C-terminal arginine was used for the calculations of average masses with the web-based tool ProtPi (`https://www.protpi.ch`). An average mass of 18,235.7324 Da was used for the EPO backbone with 2 disulfide bonds. For glycan building blocks, the residual average masses of hexose (H, 162.1406 Da), *N*-acetylhexosamine (N, 203.19252 Da), deoxyhexose (F, 146.1412 Da), *N*-acetylneuraminic acid (S, 291.2546 Da), *N*-glycolylneuraminic acid (G, 307.2544 Da), acetyl (Ac, 42.0368 Da) and phosphate (P, 79.9799 Da) were used from National Institute of Standards and Technology (`https://www.nist.gov/static/glyco-mass-calc/`). The assignments of glycoform compositions were mainly based on recent data on intact glycoforms, glycopeptides and released glycans of the studied EPO variants [71, 110, 140]. Analyte curation was based on an *m/z* tolerance of 5 Th, a signal-to-noise ratio ≥ 10, and the presence in all replicates. The assigned glycoforms were further validated by simulating the mass spectrum of doubly charged intact EPO ions, using recently published bottom-up data of EPO RS and an R script for simulating mass spectra of intact glycoproteins (`https://github.com/Yang0014/glycoNativeMS`) [59, 140]. In addition, the R script was used to assess the similarity of intact EPO RS MALDI mass spectra for the intra- and interday analysis. Relative abundances of intact EPO glycoforms were calculated by total signal intensity normalization. The data were grouped into different glycosylation traits (sialylation, acetylation and LacNAc repeats). For this, all relative intensities were summed based on the number of sialic acids, acetyl groups, or HxN(x-3) repeats. The average number of sialic acids per EPO molecule was determined by summing all relative abundances upon multiplying by the number of sialic acids.

4.3. Results and Discussion

4.3.1. Optimization of MALDI MS analysis

The analysis of sialylated glycans and glycoconjugates by MALDI MS may be affected by the loss of sialic acid residues generated from in-source and metastable decay (ISD) processes. The extent of such a loss depends on both the nature of the analyzed biomolecule and the MALDI conditions such as the MALDI matrix and the laser power. Furthermore, the ISD of biomolecules is partially influenced by the pressure of the MALDI ion source [155, 156]. It has been shown that the use of an intermediate pressure MALDI ion source minimizes the loss of sialic acid residues of released glycans and glycopeptides [157, 158]. Alternatively, sialic acid residues can be stabilized by chemical derivatization as previously shown for the analysis of released glycans and glycopeptides [159–161].

Here, different MALDI matrices, namely CHCA, ClCCA, SA, and DHAP, were evaluated for the analysis of EPO RS by MALDI FT-ICR MS (**Figure S2**). All matrices allowed for the detection of EPO RS glycoforms. The glycosylation profiles obtained using ClCCA and DHAP showed high similarity, whereas the profiles obtained with SA and CHCA showed a significant shift towards less sialylated glycoforms. This shift, presumably associated with in-source decay processes and consequent loss of sialic acid moieties, was more pronounced for the CHCA matrix. MALDI matrix adducts, detected at high intensity for ClCCA (**Figure S3**), were negligible in the mass spectra obtained using DHAP which was chosen for the further analysis of EPO RS by MALDI FT-ICR MS. The use of DHAP, as a "cold" MALDI matrix, and the intermediate pressure MALDI ion source of the FT-ICR MS system minimized sialic acid residue loss of intact EPO RS. The suitability of DHAP for the analysis of sialylated proteins has previously been reported [162].

The obtained mass spectra showed intense doubly charged EPO glycoform signals with intensities ranging over almost two orders of magnitude. The most abundant glycoforms were also detected as singly and triply charged ions but at a lower intensity (**Figure S4**). The spectra of the less complex desialylated EPO RS showed a higher abundance of singly charged ions (around 10%) and desialylated EPO RS dimers were detected as doubly and triply charged species (**Figure S4**). Due to the low number of observed charge states and their different abundance, the mass spectra were not deconvoluted, instead, the data interpretation and processing was solely based on the doubly charged ions which were measured at a resolving power of about 2,000 (**Figure S5**). Of note, MALDI FT-ICR MS has allowed the measurements of small protein and polysaccharides at an isotopic resolution up to about *m/z* 24,000, providing mass spectra of superior quality compared to MALDI TOF MS [163–175]. Ultrahigh-resolution glycosylation profiles have been obtained for apolipoprotein-CIII, ribonuclease-B, and Fc portions of monoclonal antibodies [169–173]. Isotopic resolution measurements of EPO RS did not improve the mass spectra quality (data not shown) as a consequence of the fast decay of the transient signal that affected the sensitivity of the analysis. EPO RS glycoforms, analyzed by MALDI TOF MS, were not resolved and were instead detected as a single broad peak in line with previous reports

(**Figure 4.2**) [149–152, 176]. The glycoform resolution in MALDI TOF mass spectra improved after desialylation of the EPO RS, however, also in this case, MALDI FT-ICR MS analysis was more informative.

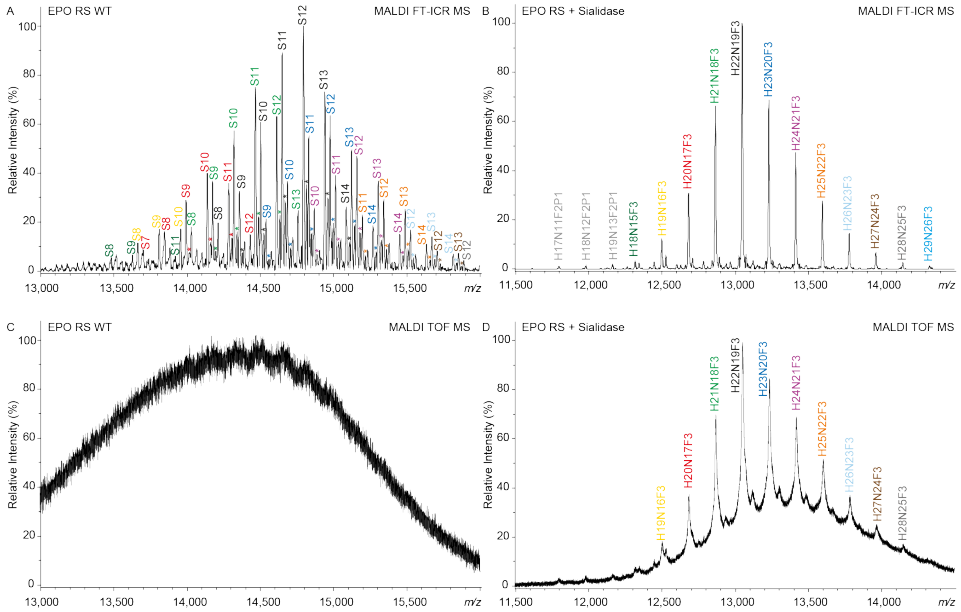

Figure 4.2: MALDI FT-ICR mass spectra of doubly charged EPO RS (**A**) wild-type, (**B**) desialylated and corresponding MALDI TOF mass spectra (**C** and **D**, respectively). *represents acetylated species.

4.3.2. GLYCOFORM ASSIGNMENTS OF EPO REFERENCE STANDARD

A complex glycosylation profile of EPO RS was obtained by MALDI FT-ICR MS (**Figure 4.2A**, **Table S1**). Desialylation of EPO RS reduced significantly the mass spectrum complexity and allowed for the assignment of 42 "core" glycoform compositions (**Figure 4.2B**, **Table S2**). All assignments were based on previously reported glycoform compositions obtained from the characterization of EPO RS samples at the intact protein, the glycopeptide, and the released glycan levels [71, 110, 140]. Thus, the main "core" glycoforms were assigned to HxN(x-3)F3 compositions generated from the combinatory occupancy of the four glycosylation sites with varying numbers of antennae and LacNAc (HN) repeats. The largest "core" glycoform was assigned to H30N27F3 while H22N19F3 had the highest abundance. A series of low-abundant HxN(x-5to6)F2P1 "core" glycoforms were also detected. These compositions were associated with phosphorylated (+ 79.9799 Da) high mannose structures (e.g. H5N2P1, H6N2P1, **Figure S6**), which are known to be present at relatively high abundances

(summed around 10%) [140]. Of note, HxN(x-6)F2P1, originating from the presence of H6N2P1 at one of the *N*-glycosylation sites, has a compositional difference of-2H-1P+2N relative to H(x-2)N(x-4)F2 resulting from an unoccupied *N*-glycosylation site which translates into a theoretical $\Delta m/z$ of -1.0620. These species could not be resolved and the detected signal was assigned to HxN(x-6)F2P1 glycoforms based on quantitative information on the high mannose glycoforms (i.e. H6N2P1, around 5%) and the occupancy of the *N*-glycosylation sites [140]. In fact, it was previously shown that Asn24 bears the highest levels of high mannose glycoforms and is the most unoccupied *N*-glycosylation site (around 2%) of the analyzed EPO RS [140]. The assignment of the "core" structures of desialylated EPO RS was further supported by the fact that only phosphorylated high mannose glycoforms were assigned in charge-sensitive high-resolution AEX-MS [71, 140]. Although deamidated proteoforms and sulfated (+ 80.063 Da) glycans of EPO RS were previously reported, these could not be resolved in our analysis [77].

The higher glycoform heterogeneity of the wild-type EPO RS is mainly derived from the presence of a varying number of sialic acids, such as Neu5Ac and to a minor extent Neu5Gc, on the "core" structures (**Figure 4.2**). Additional acetyl groups on the sialic acid residues further increased the complexity. In total, 183 glycoforms were assigned to EPO RS (**Table S1**, **Figure S7**) with the five most abundant glycoforms being H22N19F3S12, H22N19F3S11, H21N18F3S10, H22N19F3S13, H23N20F3S19, in line with previous reports on CHO cell produced EPO [59, 71]. On average, for all glycoforms present in a representative replicate, the measurement error was -0.5 ± 1.5 Th (-1.0 ± 3.0 Da) based on an external calibration, solely. Mass deviations up to 15.1 Da and 236 glycoforms were instead previously reported for intact EPO glycoforms analyzed by direct native ESI MS analysis [59]. The higher mass tolerance, the potential deconvolution artifacts and the difference in the analyzed EPO samples are possible reasons for the higher number of assigned glycoforms in ESI compared to MALDI.

The possible presence of overlapping signals generated from isobaric glycoforms or glycoforms with similar masses was also evaluated for the assignments. For example, glycoform compositions containing Neu5Gc instead of Neu5Ac differ by 15.995 Da (or $\Delta m/z$ 7.998). This mass difference can also result from the oxidation of Met54 at the protein backbone. However, no detectable amount of oxidized backbone was assigned in the desialylated EPO therefore, this protein modification was excluded. Similarly, both HxN(x-3)F3Sy and H(x+1)N(x-5)F2S(y+1)P1 ($\Delta m/z$ 9.5756) could also be assigned with high confidence. Another example is the compositional difference between 4 HN (+1461.3325 Da) and 5 NeuAc (+1456.2729 Da) that is reflected in a mass difference of 5.0596 Da ($\Delta m/z$ 2.5298). In this case, assignments were based on the mass accuracy, the assigned "core" glycoforms and previous AEX-MS assignments [71]. The assignment of acetylated glycoform compositions also required some additional consideration. Acetylation of sialic acids (up to 2 acetyl groups per sialic acid are possible) of EPO has previously been reported [11, 46, 59, 71, 110, 140]. Up to 6 acetyl groups were assigned to the most abundant glycoform (H22N19F3S12) of EPO RS (**Table S1**). Glycoforms with one acetyl group do not overlap with other glycoforms and can be accurately assigned

whereas glycoforms with two acetyl groups (HxN(x-3)F3SyAc2) were rarely assigned since glycoforms containing H(x+1)N(x-2)F3S(y-2)G1, with similar masses ($\Delta m/z$ 3.0024), were generally more prominent based on mass accuracy. For compositions with a higher number of acetyl groups, another partial overlap ($\Delta m/z$ 4.9975) between HxN(x-3)F3SyAcz and H(x+1)N(x-2)F3S(y-1)Ac(z-2) was in addition carefully considered for more confident assignments.

Considering the degree of complexity, structural and compositional information obtained using other types of analysis is pivotal to make confident assignments in MALDI FT-ICR mass spectra. In addition to the reported glycosylation analyses of EPO RS, the assignments were further supported by a comparison between the MALDI FT-ICR MS spectra of EPO RS and an in silico constructed spectrum obtained from bottom-up glycopeptide data (**Table S3**) [140], as previously reported by Yang and coworkers [59]. Such a constructed glycosylation profile of EPO RS showed high similarity to the glycoform distribution in the acquired MALDI mass spectrum (**Figure S8**) although a shift towards higher sialylation levels was observed in the measured data. This was in line with what was previously shown by Yang and coworkers which compared ESI MS spectra of intact EPO and an in silico constructed spectra from glycopeptides data [59]. The observed differences may be explained by the loss of sialic acids and an ionization bias of the sialic acids in bottom-up glycopeptide analysis compared to intact protein characterization. Finally, the glycosylation profiles obtained with MALDI FT-ICR MS were in agreement with previously reported data obtained using bottom-up and intact protein analysis by ESI MS [71, 140].

To assess the repeatability of the EPO RS glycosylation profiles obtained by MALDI FT-ICR MS, replicate measurements of EPO RS were performed over three days. The median intraday and interday variabilities, calculated for the relative intensities of all glycoform compositions (n=183), were 7.4% and 9.8%, respectively (**Table S4**). The good repeatability of the EPO glycosylation profiles was corroborated by mass spectra similarity scores higher than 0.96 (**Figure S9**).

4.3.3. GLYCOSYLATION ANALYSIS OF EPO VARIANTS

Detailed glycosylation profiles of two additional EPO variants, namely EPO RP+ and EPO RP-, were obtained using the MALDI FT-ICR MS method optimized for the analysis of EPO RS (**Figure 4.3**, **Tables S5** and **S6**). These variants were recently characterized by QC release methods, bottom-up LC-MS and AEX-MS at the intact level [71, 122, 140]. In the MALDI FT-ICR MS spectra, 169 and 199 glycoform compositions were assigned to EPO RP+ and EPO RP-, respectively, whereas 160 and 192 glycovariants (deamidated proteoforms not included) were previously assigned by AEX-MS. A higher number of LacNAc repeats was detected by MALDI FT-ICR MS for EPO RP+ compared to EPO RS, in line with bottom-up LC-MS data (mainly at Asn83) [122]. The MALDI spectra of EPO RP- indicated lower levels of *N*-glycosylation. Instead, glycoforms with only two occupied *N*-glycosylation sites (HxNx-2F2Sy) were assigned for EPO RP-, e.g. the most

abundant species was assigned to H15N13F2S9 instead of H17N11F2S9P1 (assigned glycoform for EPO RS), in line with the previously reported low site occupancy of Asn24 [122].

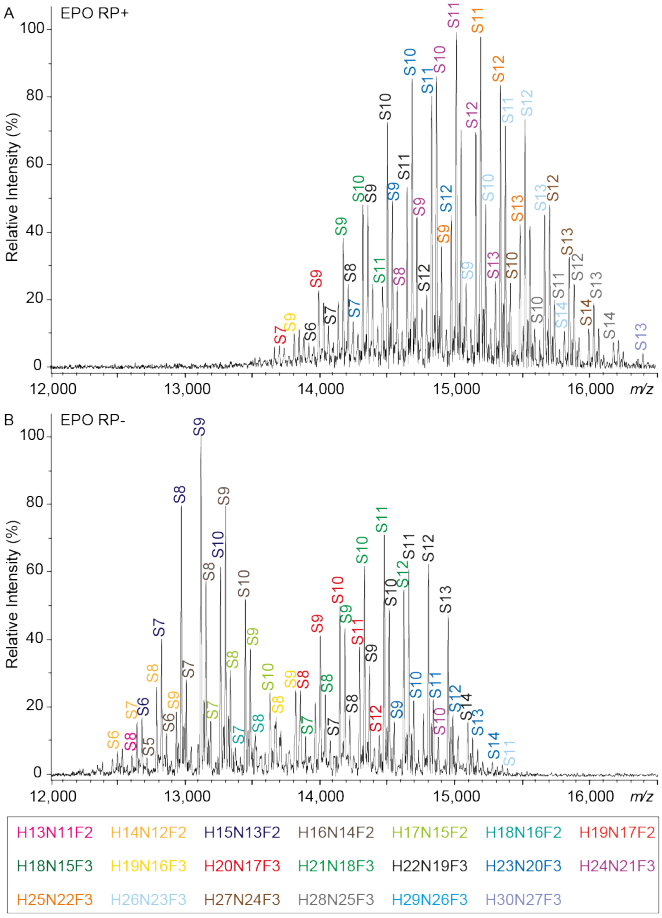

Figure 4.3: MALDI-FT-ICR mass spectra of EPO variants with (**A**) increased number of LacNAc units and (**B**) decreased glycosylation site occupancy.

Key glycosylation features of the three different EPO variants were compared (**Figure 4.4**). The average number of sialic acids, determined by MALDI-FT-ICR MS, was 10.7 (EPO RS), 10.9 (EPO RP+) and 9.4 (EPO RP-). This is in good agreement with QC release, bottom-up LC-MS and AEX-MS analyses previously reported for the same EPO variants (**Table 4.1**). MALDI FT-ICR MS showed the same sialylation ranking (EPO RP+ > EPO RS > EPO RP-), except for the AEX-MS method (EPO RS > EPO RP+ > EPO RP-). Considering the QC release method as the best estimation of sialylation, the sialic acid stability of the MALDI FT-ICR MS method is above 95%.

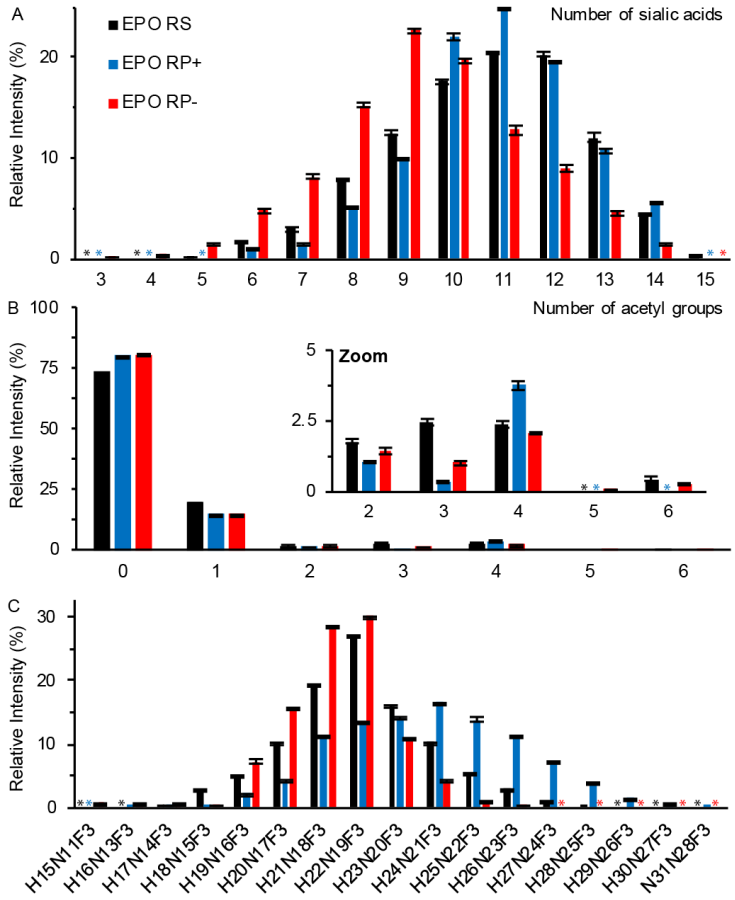

Figure 4.4: Key glycosylation features of three EPO variants analyzed by MALDI FT-ICR MS. (**A**) sialylation, (**B**) acetylation and (**C**) LacNAc repeats. * not detected.

Table 4.1: Comparison of the average sialic acid content per EPO molecule determined by different methods.

Method	EPO RS	EPO RP+	EPO RP-
MALDI FT-ICR MS	10.7 ± 0.03	10.9 ± 0.01	9.4 ± 0.04
QC* [122]	11.2 ± 0.05	11.4***	10.2***
MAM** [122]	10.9 ± 0.1	10.4***	10.1***
AEX-MS [71]	11.6 ± 0.04	11.4 ± 0.08	11.0 ± 0.13

* HPAEC-PAD analysis of released sialic acids (Neu5Ac).
** Multiple Attribute Monitoring by bottom-up LC-MS of a GluC digest.
*** Standard deviations not available.

Other glycosylation quality features such as LacNAc repeats and acetylation can also be derived from the MALDI FT-ICR mass spectra (**Figure 4.4**). In line with previous reports, 70% - 80% of the sialic acids are not acetylated [71]. Mono-acetylated species were below 20% and higher degrees of acetylated glycoforms should be carefully interpreted owing to described overlapping species. The expected differences in the LacNAc repeats of the different EPO variants could be shown as well by the relative quantification (**Figure 4.4**). The distribution of HxN(x-3)F3 species obtained for EPO RS was in good agreement with the analysis upon desialylation (**Figure 4.2**). EPO RP+ showed the highest content of HN units (H24N21F3 most abundant) and both EPO RS and EPO RP- showed H22N19F3 as the most abundant species.

It should be stressed that prior knowledge of glycosylation moieties of complex glycoproteins, such as EPO, is needed for a reliable interpretation of intact protein mass spectra. Such information must be obtained by a multi-level structural characterization of the glycoproteins which includes released glycan analysis and site-specific characterization by bottom-up MS. Nonetheless, as depicted in **Figure 4.1**, a comprehensive characterization of EPO must include the analysis at the intact protein level to determine the combinatory presence of protein modifications on different modification sites. This unique information is key for assessing the structural integrity, quality, and safety of pharmaceutical proteins.

4.4. CONCLUSIONS

In this study, we showed that detailed glycosylation profiles of intact EPO can be obtained by low-resolution MALDI FT-ICR MS. An exceptional improvement in glycoform resolution was achieved compared to MALDI TOF MS analysis. MALDI FT-ICR MS provided glycosylation profiles comparable to data previously obtained with other types of analyses such as native ESI MS and AEX-MS and allowed for a more straightforward data acquisition and interpretation. In fact, MALDI FT-ICR MS allowed for the measurements of multiple samples without the need for in-between washing steps as for native ESI MS, while the evaluation of only doubly charged ions made deconvolution unnecessary. The comparison of the glycosylation profiles of three different EPO variants, namely EPO RS, EPO RP+ and EPO RP-, showed that MALDI FT-ICR MS allows for the evaluation of key glycosylation features such as the number of sialic acid residues and acetyl groups and LacNAc repeats. In particular, sialylation – the most important and susceptible glycosylation trait of EPO – was proven to be stable under the chosen MALDI conditions and sialylation profiles were in good agreement with those obtained from other orthogonal methods. Of note, as the loss of sialic acid during MALDI also depends on the nature of the analyzed biomolecule, different spotting conditions and/or acquisition parameters may be required for the analysis of different glycoproteins. Finally, the simplicity and potential high-throughput capability of the here developed method are desirable characteristics for applications in the biopharmaceutical industry. Further studies are needed to evaluate the applicability of the developed method for the analysis of unpurified EPO during the manufacturing

process and to evaluate (and develop) more accessible MS technology (e.g. MALDI qTOF MS) that could provide similar mass spectral quality of EPO glycoforms as obtained using 15T MALDI FT-ICR MS. To our knowledge, this is the first report on the use of MALDI MS for the glycoform-resolved analysis of highly sialylated intact glycoproteins. We anticipate the application of our strategy for the analysis of different pharmaceutical glycoproteins with a molecular mass of up to 40 kDa including various hormones and glycoconjugate vaccines.

SUPPORTING INFORMATION

Supporting information is available free of charge via https://ars.els-cdn.com/content/image/1-s2.0-S0003267021009107-mmc1.docx and https://ars.els-cdn.com/content/image/1-s2.0-S0003267021009107-mmc2.xlsx.

ACKNOWLEDGEMENTS

This research was funded by the European Commission H2020 (Analytics for Biologics project, Grant 765502).

5

GLYCOFORM-RESOLVED FCγRIIIA AFFINITY CHROMATOGRAPHY–MASS SPECTROMETRY

Steffen Lippold, [1] Simone Nicolardi, [1] Elena Domínguez-Vega, [1] Anna-Katharina Heidenreich, [2] Gestur Vidarsson, [3] Dietmar Reusch, [2] Markus Haberger,[2] Manfred Wuhrer[1] and David Falck[1]

[1]Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands
[2]Pharma Technical Development Penzberg, Roche Diagnostics GmbH, Penzberg, Germany
[3]Department of Experimental Immunohematology, Sanquin Research and Landsteiner Laboratory, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

Reprinted and adapted from mAbs, 11:7, 1191-1196
DOI: 10.1080/19420862.2019.1636602

Determination of the impact of individual antibody glycoforms on FcγRIIIa affinity, and consequently antibody-dependent cell-mediated cytotoxicity (ADCC) previously required high purity glycoengineering. We hyphenated FcγRIIIa affinity chromatography to mass spectrometry, which allowed direct affinity comparison of glycoforms of intact monoclonal antibodies. The approach enabled reproduction and refinement of known glycosylation effects, and insights on afucosylation pairing as well as on low-abundant, unstudied glycoforms. Our method greatly improves the understanding of individual glycoform structure-function relationships. Thus, it is highly relevant for assessing Fc-glycosylation critical quality attributes related to ADCC.

Fcγ receptors (FcγR) are key elements in many immunological responses. FcγRIIIa has an important role in mediating antibody-dependent cell-mediated cytotoxicity (ADCC). Thus, investigating antibody-FcγRIIIa interactions is highly relevant for fundamental and clinical research, as well as (bio)therapeutic innovations [32, 177]. The Fc glycosylation of immunoglobulin G (IgG) has large effects on its FcγRIIIa interaction [33]. Most significantly, IgG fucosylation reduces FcγRIIIa affinity by up to 100-fold. This is in part attributable to a unique glycan-glycan interaction between the IgG Fc and the FcγRIIIa N162 glycan [31]. Knowledge of the effects of IgG glycosylation led to the successful development and approval of next-generation, glycoengineered therapeutic monoclonal antibodies (mAbs) with low fucose levels, such as Gazyva (obinutuzumab, Roche/Genentech) [37, 178]. In contrast to the effects of fucosylation, galactosylation of the IgG Fc glycan slightly increases the FcγRIIIa affinity [34, 35]. Several glycosylation traits are therefore critical quality attributes [12].

Analytical methods for assessing effector functions and Fc/Fc-receptor interactions in therapeutic antibodies were recently reviewed [177, 179]. Functional cellular activity assays have the advantage of high relevance towards the *in vivo* situation. The downstream signaling is determined by a complex interplay of immune complexes binding to activating and inhibitory FcγR [32]. The interaction of FcγRIIIa with the Fc portion occurs in an asymmetrical 1:1 stoichiometry that precludes cellular activation *in vivo* by monomeric antibodies [32, 180]. Thus, the translation of functional cellular assays is more straightforward compared to cellular or cell-free binding assays studying monomeric antibodies. However, the fundamental interactions of Fc and FcγRIIIa do not depend on whether IgG is present in monomeric form or as an immune complex, and it is important to study them. Several cell-free physicochemical assays are well established and generally, the first choice for assessing the potential impact of monomeric IgG Fc attributes on receptor binding due to their increased analytical performance with regard to assay complexity, affinity resolution, and robustness [179]. Moreover, these binding assays strongly linked *in vitro* FcγRIIIa affinity and ADCC activity [8, 34, 35]. Explicitly, retention time differences in FcγRIIIa affinity liquid chromatography (AC) were linked to ADCC [34].

A common disadvantage of all previous methods is the averaged and potentially biased output because of the naturally occurring IgG glycoform heterogeneity. Therefore, unraveling the impact of individual IgG glycoforms on FcγRIIIa affinity previously required laborious glycoengineering. The lack of molecular resolution of established *in vitro* affinity assessment techniques, such as surface plasmon resonance (SPR) or AC necessitated high IgG glycoform purity [8, 34, 181].

Here, we present the simultaneous assessment of FcγRIIIa affinity of multiple IgG glycoforms of a therapeutic mAb. This was achieved by AC hyphenated to mass spectrometry (MS). MS allowed molecular resolution while the separation dimension provides FcγRIIIa (V158) affinity. Advantageous features are: (1) affinity assessment of individual, previously unstudied glycoforms from biosynthetic mixtures, omitting the need for advanced glycoengineering; and (2) increased affinity differentiation

compared to established *in vitro* techniques due to the simultaneous assessment.

MS has become an important technique for characterizing intact proteins [182–185]. The combination with AC (AC-MS) proved its potential for functional characterization of therapeutic mAbs recently. AC-MS based on the fetal/neonatal Fc receptor (FcRn) showed decreased FcRn affinity, and by extension IgG half-life, for a mAb oxidized at M255 [72]. However, FcRn affinity is only very weakly influenced by glycosylation [8, 181]. Therefore, resolving complex glycosylation heterogeneity molecularly on an intact protein level is necessary for FcγRIIIa-AC-MS, which makes it more challenging and more powerful at the same time. For method development, a good balance between MS response and separation efficiency is crucial. In contrast to previous AC-UV studies [181], we employed a simple ammonium acetate buffer. We optimized ammonium acetate concentration and linear pH gradient (**Supplemental Figure 1**), aiming at an at least equal separation efficiency compared to the previously reported FcγRIIIa AC-UV conditions (**Figure 5.1**(**a**)). This was achieved using 50 mM ammonium acetate and a pH gradient from pH 5 - pH 3 (**Figure 5.1**(**b**)). We used the same column for which FcγRIIIa AC retention times were previously linked to ADCC and obtained comparable profiles (**Figure 5.1**) [34].

A 15 T Fourier transform ion cyclotron resonance (FT-ICR)-MS instrument provided high sensitivity and mass accuracy for determining even low-abundant glycoforms (**Figure 5.1**(**c,d**), **Supplemental Table 3** and **4**). We detected 21 compositions, reflecting 27 partially isomeric glycoforms (**Figure 5.1**(**b**) and **Supplemental Table 4**). Typical charge state distributions ($[M + 24H]^{24+}$ to $[M + 29H]^{29+}$) of folded protein conformations were observed for the entire pH gradient (inserts **Figure 5.1**(**c,d**)). The method showed very good intra- and inter-day variability of relative peak areas for the resolved glycoforms of the mAb (**Supplemental Figure 3** and **Supplemental Table 5**). Further, relative peak areas of glycoforms of the intact mAb showed good agreement compared to the glycoform profile based on released glycan analysis (**Supplemental Figure 5**). This demonstrated the feasibility of performing relative quantitation and provides additional proof of correct glycoform assignment.

AC-MS enables the simultaneous *in vitro* comparison of FcγRIIIa affinity of antibody glycoform mixtures, in contrast to SPR or AC-UV. Previously, FcγRIIIa affinities of seven glycoengineered mAbs were assessed using AC-UV [181]. More recently, AC-UV and subsequent glycan analysis were used to assess the heterogeneity of Fc-glycosylation of a therapeutic antibody, using a non-glycosylated FcγRIIIa [186]. However, glycosylation has been shown to be critical for differential recognition of IgG-glycoforms by wild-type FcγRIIIa [187]. Thus, no significant differences for individual glycoforms, except for galactosylation levels, were observed. With our method, we can directly compare the affinity of 21 glycoforms present in a therapeutic mAb without the need for fractionation or glycoengineering. Herein, we use a human embryonic kidney (HEK) cell-produced FcγRIIIa. This may lead to altered glycosylation and affect affinity/selectivity compared to the natural receptor. Currently, available natural FcγRIIIa data is missing site-specificity or is inferred (and incomplete) [187, 188].

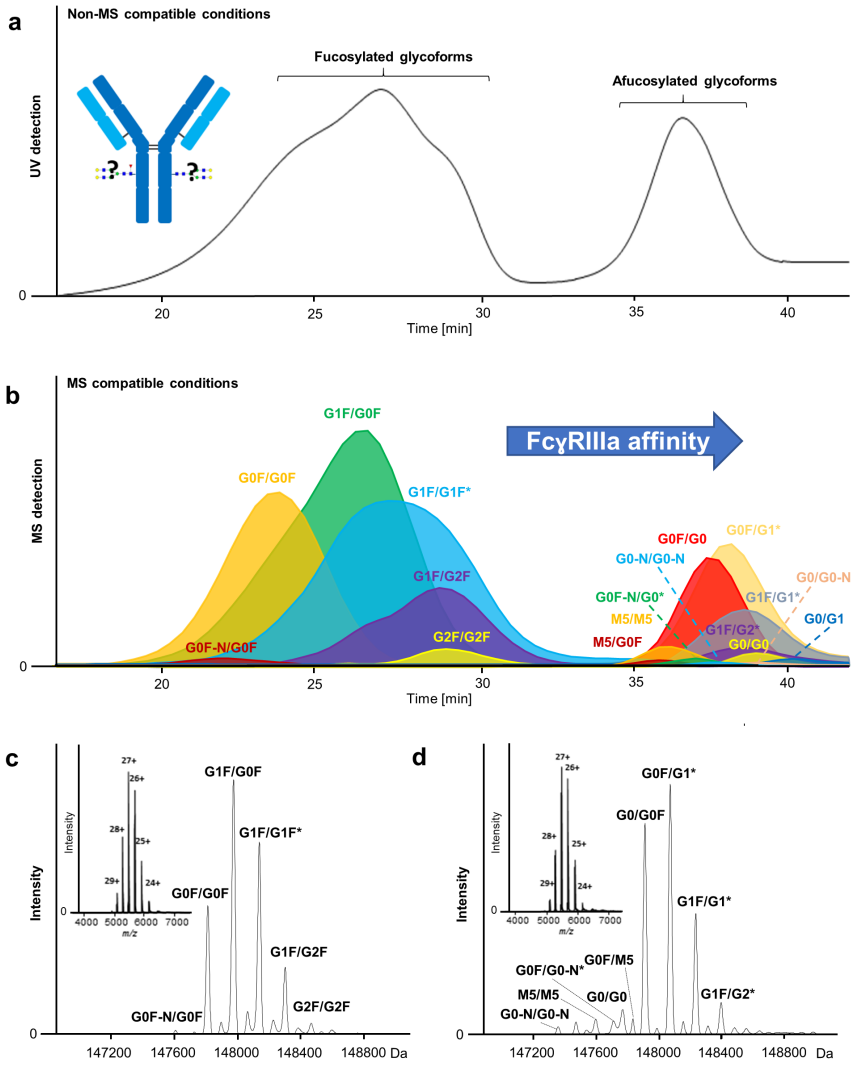

Figure 5.1: FcγRIIIa affinity chromatography for a therapeutic mAb. (**a**) UV chromatogram using reported non-MS compatible conditions [181]. (**b**) AC-MS under MS-compatible conditions represented by extracted ion chromatograms of detected glycoforms. (**c** and **d**) Deconvoluted mass spectra and charge state distribution (inserts) of (c) 2x fucosylated (18 - 34 min) and (d) remaining glycoforms (34 - 42 min). In case of multiple possibilities (asterisk), the most probable glycoform is presented, based on reference data (**Supplemental Table 7**); all possible glycoforms and their structures are listed in **Supplemental Table 1** and **4**.

While the majority of CHO and HEK cell glycoforms seems to qualitatively overlap with natural glycosylation, neither seems good at producing triantennary structures [189]. HEK cells can produce antennary fucosylation, but also unnatural LacDiNAc structures [189]. At the same time, the functional impact of individual FcγRIIIa glycoforms is unclear.

Fucosylation is known to have the highest impact on FcγRIIIa affinity [29, 33]. Single afucosylation is sufficient to drastically increase binding, owing to the asymmetric 1:1 binding between IgG and FcγRIIIa [28, 180]. Consequently, glycoforms containing fucosylation on both heavy chains (2x fucosylated) populate the earlier eluting peak in the AC chromatogram (**Figure 5.1**(**b**,**c**)). The 1x (a)fucosylated and 2x afucosylated species are found in the later eluting AC peak (**Figure 5.1**(**b**,**d**)). However, the AC-MS system achieves a much more detailed differentiation of affinity than AC-UV, because it can rank FcγRIIIa affinity of multiple (partially) co-eluting glycoforms in one run (**Supplemental Table 6**). In contrast, AC-UV requires comparison of multiple runs, making retention time stability a limiting factor. Other assays only give an average value for all glycoforms present, resulting in a potentially significant impact of even low-abundant glycoforms of vastly higher affinity. As a result, we could demonstrate that 2x afucosylated complex glycoforms (G0(N)/G0(N), G0(N)/G1(N)) show an increased FcγRIIIa affinity over 1x (a)fucosylated (**Figures 5.1**(**b**) and **5.2**(**a**)). Furthermore, AC-MS can distinguish the positive influence of galactosylation on FcγRIIIa affinity in all fucosylation variants (**Figures 5.1**(**b**), **5.2**(**a**), **Supplemental Figure 4**). Previously, this was only observed for 2x fucosylated glycoforms by AC-UV and mainly 1x afucosylated glycoforms by AC-UV and SPR [34, 35, 181]. More recently, a comprehensive binding study applying AC-UV, SPR and ADCC analysis showed the positive influence of afucosylation and galactosylation for highly homogeneous 2x fucosylated glycoforms and 0x fucosylated glycoforms [8].

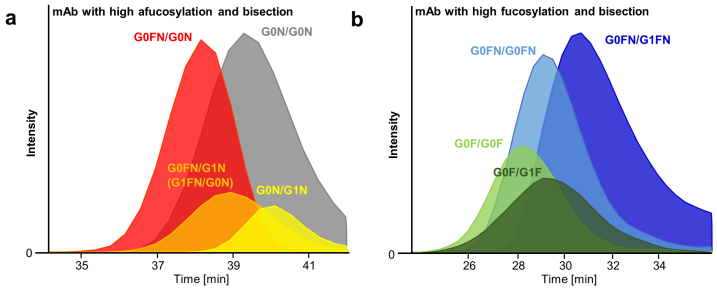

Figure 5.2: Extracted ion chromatograms of AC-MS with glycoengineered mAbs containing a high level of bisected and afucosylated glycoforms (**a**) and high level of bisected and fucosylated glycoforms (**b**). Increased affinity is observed for afucosylation compared to 1x (a)fucosylation. Bisected fucosylated glycoforms showed also increased affinity compared to non-bisected fucosylated glycoforms. Positive effect of galactosylation is represented for each glycosylation feature.

A further challenge in glycoengineering is biosynthetic interferences between the different glycosylation features [35, 190]. For example, bisection correlates with other glycosylation features depending on the expression system, which complicates independent affinity assessment [190, 191]. Furthermore, glycoengineered mAbs often still exhibit a high degree of glycosylation heterogeneity [191]. The high molecular resolution of our method allows the study of subtle FcγRIIIa affinity differences by dissecting the contribution of individual glycosylation features. For example, we could demonstrate that bisection increases FcγRIIIa affinity independent of afucosylation and galactosylation [35, 191]. This was observed regardless of the presence of galactoses (**Figure 5.2**(**b**)).

In addition, FcγRIIIa affinity of typically low-abundant glycoforms (≈0.1% relative abundance, **Supplemental Table 5**) could be studied (**Figure 5.1**(**b**)). For example, mono-antennary glycoforms lacking N-acetylglucosamine (-N) showed decreased affinity compared to its complex bi-antennary glycoforms (e.g., G0F/G0F-N vs. G0F/G0F, **Figure 5.1**(**b**), **Supplemental Figure 4** and **Supplemental Table 6**). Glycoengineering, and therefore classical affinity assessment, of low-abundant species, is often deemed too tedious, a limitation efficiently addressed by our method.

High mannose species have been shown to increase ADCC activity via FcγRIIIa binding that was attributed to the lack of core-fucose [192]. A minor amount of high mannose glycans, paired with either another high mannose (0.8% M5/M5) or the complex glycan (0.2% M5/G0F), was detected in the therapeutic mAb (**Figure 5.1**(**b**)) [193]. Largely, these showed higher affinity compared to 2x fucosylated glycoforms. However, the affinity of M5-containing glycoforms was reduced compared to 1x (a)fucosylated complex glycoforms. These findings are supported by previous studies describing increased binding affinities [8, 181, 192] and ADCC activities [8, 192] of high mannose glycoforms over 2x fucosylated complex glycoforms, and similar or decreased affinities and activities compared to afucosylated complex glycoforms [8, 181, 192]. Identity and affinity of the M5/M5 were confirmed by glycoengineering (**Supplemental Figure 6 and 7**). Moreover, the isolated analysis of the glycoengineered M5/M5 mAb showed comparable retention as observed in the therapeutic mAb, giving no indication for a measurable bias in retention behavior caused by competitive receptor binding of other glycoforms.

Glycoforms with one non-glycosylated heavy chain (X/NG) are normally also found to a low extent in therapeutic antibodies [194]. Decreased binding affinities and ADCC activities were observed for these glycoforms in previous studies [195]. We observed elution of fucosylated glycoforms (G0F/NG, G1F/NG, G2F/NG) already before the pH gradient (for initial pH of 5) started, whereas G0/NG showed the lowest affinity of the afucosylated glycoforms, but it exhibited a higher affinity compared to the 2x fucosylated glycoforms (**Supplemental Table 6**, **Supplemental Figure 4**).

In conclusion, FcγRIIIa AC-MS represents a great advance in characterizing the impact of antibody Fc glycosylation on FcγRIIIa affinity. It omits the need for high purity

glycoengineering, which is notoriously difficult, especially for biosynthetic pathway intermediates [28], and has hitherto complicated and, in part, prohibited the affinity assessment of various glycoforms. We were able to perform an FcγRIIIa affinity assessment of individual glycoform pairings present in a therapeutic mAb, including so far unstudied, low-abundant glycoforms (e.g., G0F-N/G0F). Due to its high molecular and high FcγRIIIa affinity resolution, the method allows independent assessment of individual glycosylation features. For example, we could show, for the first time, the increased affinity of afucosylated complex glycoforms over singly (a)fucosylated complex glycoforms and high mannose glycoforms. Furthermore, we could support the independent, positive influence of bisection. FcγRIIIa AC-MS will allow a fast and reliable assessment of FcγRIIIa affinity of novel (synthetic) glycoforms, even in the presence or excess of other glycoforms.

FcγRIIIa AC-MS may be highly useful in other applications. For example, other post-translational modifications as well as protein backbone differences may be analyzed in addition to glycosylation, again separating independent effects. Thus, FcγRIIIa AC-MS will be an enormously valuable tool in studying and tuning ADCC activation of (therapeutic) antibodies. Largely by the omission of high purity glycoengineering, proteoform-related critical quality attributes and their interactions could be defined earlier in therapeutic mAb development. In the future, glycoengineering of FcγRIIIa towards natural glycosylation would be highly desirable. In general, AC-MS approaches have a high potential for studying biomolecular interactions in a proteoform-resolved manner and are extremely valuable in unraveling structure-function relationships.

5.1. MATERIALS AND METHODS

5.1.1. CHEMICALS

All chemicals were purchased from Sigma-Aldrich (Steinheim, Germany) and had at least analytical grade quality. Deionized water was used from a Purelab ultra (ELGA Labwater, Ede, The Netherlands). Ammonium acetate solution (7.5 M, Sigma-Aldrich) and glacial acetic acid (Fluka-Honeywell) were used for preparing MS-compatible mobile phases.

5.1.2. ANTIBODIES

In this study, four different monoclonal antibodies (mAbs) differing in glycosylation were used. First, a standard therapeutic mAb covering a typical degree of glycosylation heterogeneity was investigated. In addition, a glycoengineered high mannose-containing variant of the first mAb and a mAb contained a high level of afucosylation and bisection were analyzed. Lastly, a highly fucosylated and highly bisected anti-D mAb were investigated (reported as +B) [35]. The first three mAbs were

provided by Roche Diagnostics (Penzberg, Germany) and the last one by Sanquin (Amsterdam, The Netherlands).

5.1.3. FcγRIIIA COLUMN PREPARATION

FcγRIIIa affinity column was provided by Roche Diagnostics (Penzberg, Germany). The preparation of the FcγRIIIa affinity column is reported elsewhere.8 In brief, biotinylated human FcγRIIIa_V158 (HEK cell-produced with glycosylation profile as reported [189]) was incubated with streptavidin sepharose for 2 h upon mild shaking and the receptor-derivatized sepharose was packed in a Tricorn 5/50 Column housing (inner diameter 5 mm x length 50 mm, GE Healthcare).

5.1.4. FcγRIIIA AFFINITY LIQUID CHROMATOGRAPHY

For FcγRIIIa affinity liquid chromatography with non-MS compatible conditions and UV detection (absorbance measured at 280 nm), previously reported pH gradient and mobile phase composition were used.13 Equilibration buffer/mobile Phase A consisted of 20 mM sodium citrate pH 6.0 and 150 mM NaCl. Mobile phase B contained 20 mM citrate pH 3.0 and 150 mM NaCl. Final MS-compatible conditions contained 50 mM ammonium acetate pH 5.0 (mobile phase A) and 50 mM ammonium acetate pH 3.0 (mobile phase B). All experiments were performed using a biocompatible Thermo Ultimate3000 instrument at 25°C. The samples were diluted with mobile phase A, if necessary (conc. 1-5 µg/µL). Injection volume was set to 10 µL. The flow rate was set to 500 µL/min. For hyphenation with MS, the flow rate was reduced from 500 µL/min to 30 µL/min via flow-splitting prior to the electrospray ionization (ESI) source. Upon injection, the column was washed with 5 column volumes (10.5 min) of mobile phase A, and the sample was eluted in a linear gradient of 15 column volumes (31.5 min) ending with 100% mobile phase B. In addition, a washing step using 5 column volumes (10.5 min) of mobile phase B and a linear pH gradient of 7.5 min ending with 100% mobile phase A was applied after the elution gradient. Re-equilibration using 15 column volumes (31.5 min) of mobile phase A was applied between runs.

5.1.5. FOURIER TRANSFORM ION CYCLOTRON RESONANCE MASS SPECTROMETRY

Online ESI-MS coupling of FcγRIIIa AC was performed on a Bruker 15 T solariX FT-ICR-MS (Bruker Daltonics, Bremen, Germany). Instrument tuning and calibration were performed using direct infusion (2 µL/min) of the therapeutic mAb (0.1 µg/µL in mobile phase A) (**Supplemental Figure 2** and **Table 2**). The *m/z* of all charge states of the main glycoform (G0F/G1F) were used for calibration. By using the charge state envelope of the most abundant glycoform, the relevant mass range was sufficiently covered. External calibration was not needed for the intended purpose of the study

because the masses and glycoforms of the analyzed mAbs were known. ESI capillary voltage was set to 4,000 V and endplate offset to -500 V. Nebulizer gas pressure was set to 0.8 bar, dry gas to 3 L/min and source temperature to 200°C. The ion funnel 1 was operated at 150 V, radio frequency amplitude at 300 Vpp and skimmer 1 at 125 V. Spectra were acquired in an *m/z*-range of 506-20,000. Accumulation time was set to 1 s. The amount of data points was set to 128 k. Each spectrum in serial mode analysis resulted from the summation of 20 spectra. This resulted in the acquisition of 2.6 data points per minute in the chromatogram. This acquisition rate was chosen as a compromise of sensitivity (for low-abundant glycoforms) and resolution, and it was sufficient for the generally broad eluting peaks observed in this study. All mass spectra were visualized and processed using DataAnalysis 5.0 (Bruker Daltonics). For deconvolution, the Maximum Entropy tool was used (145,000-150,000 Da, datapoint spacing = 1, instrument resolving power = 3,000). Further information regarding data processing are described in the **Supplemental Material**.

5 SUPPORTING INFORMATION

Supporting information is available free of charge via `https://www.tandfonline.com/doi/suppl/10.1080/19420862.2019.1636602/suppl_file/kmab_a_1636602_sm1378.docx`.

ACKNOWLEDGEMENTS

This research was supported by Roche Diagnostics, the European Union (Analytics for biologics project, Grant No. 765502), and the NWO (Vernieuwingsimpuls Veni Project No. 722.016.008 and SATIN, Project No. 731.017.202). We are grateful to Tilman Schlothauer from Roche Diagnostics for providing the FcγRIIIa column.

6

Proteoform-Resolved FcγRIIIa Binding Assay for Fab Glycosylated Monoclonal Antibodies Achieved by Affinity Chromatography Mass Spectrometry of Fc Moieties

Steffen Lippold, [1] Simone Nicolardi, [1] Manfred Wuhrer[1] and David Falck[1]

[1] *Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands*

Reprinted and adapted from Front. Chem.
DOI: 10.3389/fchem.2019.00698

Fcγ receptors (FcγR) mediate key functions in immunological responses. For instance, FcγRIIIa is involved in antibody-dependent cell-mediated cytotoxicity (ADCC). FcγRIIIa interacts with the fragment crystallizable (Fc) of immunoglobulin G (IgG). This interaction is known to be highly dependent on IgG Fc glycosylation. Thus, the impact of glycosylation features on this interaction has been investigated in several studies by numerous analytical and biochemical techniques. FcγRIIIa affinity chromatography (AC) hyphenated to mass spectrometry (MS) is a powerful tool to address co-occurring Fc glycosylation heterogeneity of monoclonal antibodies (mAbs). However, MS analysis of mAbs at the intact level may provide limited proteoform resolution, for example, when additional heterogeneity is present, such as antigen-binding fragment (Fab) glycosylation. Therefore, we investigated middle-up approaches to remove the Fab and performed AC-MS on the IgG Fc to evaluate its utility for FcγRIIIa affinity assessment compared to intact IgG analysis. We found the protease Kgp to be particularly suitable for a middle-up FcγRIIIa AC-MS workflow as demonstrated for the Fab glycosylated cetuximab. The complexity of the mass spectra of Kgp digested cetuximab was significantly reduced compared to the intact level while affinity was fully retained. This enabled a reliable assignment and relative quantitation of Fc glycoforms in FcγRIIIa AC-MS. In conclusion, our workflow allows a functional separation of differentially glycosylated IgG Fc. Consequently, applicability of FcγRIIIa AC-MS is extended to Fab glycosylated IgG, i.e., cetuximab, by significantly reducing ambiguities in glycoform assignment vs. intact analysis.

6.1. INTRODUCTION

THE fragment crystallizable (Fc) of antibodies mediates immunological responses, for example through binding to Fc receptors [32, 196]. Fc glycosylation has a key role in modulating Fc receptor-mediated effector functions, such as antibody-dependent cell-mediated cytotoxicity (ADCC) [12, 179, 197]. The affinity toward FcγRIIIa is known to be crucial for ADCC [32]. Fucosylation of Fc glycans drastically decreases FcγRIIIa affinity which is attributable to an unique glycan-glycan interaction [31]. Other glycosylation features such as galactosylation were also shown to affect the Fc-FcγRIIIa interaction [34, 35]. The binding of the Fc to FcγRIIIa is asymmetric in a 1:1 stoichiometry [180]. Nonetheless, FcγRIIIa affinity is influenced by the pairing of Fc glycans [28]. While differential affinity of glycoforms is dominated by the stronger binding glycan, the second glycan modulates affinity to a smaller extent, but along the same structural features [28, 198]. Nowadays, therapeutic monoclonal antibodies (mAbs) are most often derived from human immunoglobulin G 1 (IgG1, schematic overview in **Figure 6.1**). They are used in the treatment of various diseases, such as cancers or autoimmune diseases [199, 200]. In the biopharmaceutical industry, mAbs are very successful and currently dominate new approvals [1]. Recently, glycoengineering for enhanced FcγRIIIa affinity and ADCC has been therapeutically exploited [29, 37].

Numerous analytical technologies exist for assessing the effector functions of therapeutic antibodies [177, 179]. They vary largely in information content, generally with a negative correlation between complexity and resolution. Complex cellular assays are more easily translated to the *in vivo* situation. Contrary, physicochemical assays provide higher molecular resolution and better robustness. Though immune responses depend on the formation of immune complexes, receptor binding studies on monomeric IgG are highly relevant and widely used [32, 179] Ultimately, combining information from different assays is essential to fully understand antibody effector functions. Glycosylation heterogeneity is a major challenge for the assessment of individual contributions of specific glycoforms to the effector functions, especially considering pairing possibilities. Several studies applied laborious glycoengineering in order to assess receptor binding and effector functions of specific glycoforms [8, 34, 35, 181]. Affinity chromatography (AC) represents a cell-free physicochemical assay which provides a functional separation and correlates well with surface plasmon resonance (SPR) assays and ADCC assays [8, 34, 181]. We reported recently on coupling of FcγRIIIa AC to mass spectrometry (AC-MS) [198]. This approach allows the differential assessment of Fc glycoforms in heterogeneously glycosylated mAbs with high resolution of proteoforms and affinity on an intact protein level. Whereas it should be very powerful for most mAbs, proteoform resolution may be insufficient for more complex formats [61]. This applies to mAbs with a higher degree of heterogeneity due to sequence variants or post translational modifications (PTMs), especially additional glycosylation sites in the antigen-binding fragment (Fab). In addition, the analysis of new antibody-derived therapeutic formats, such as bispecific antibodies or fusion

proteins, may be challenging [201].

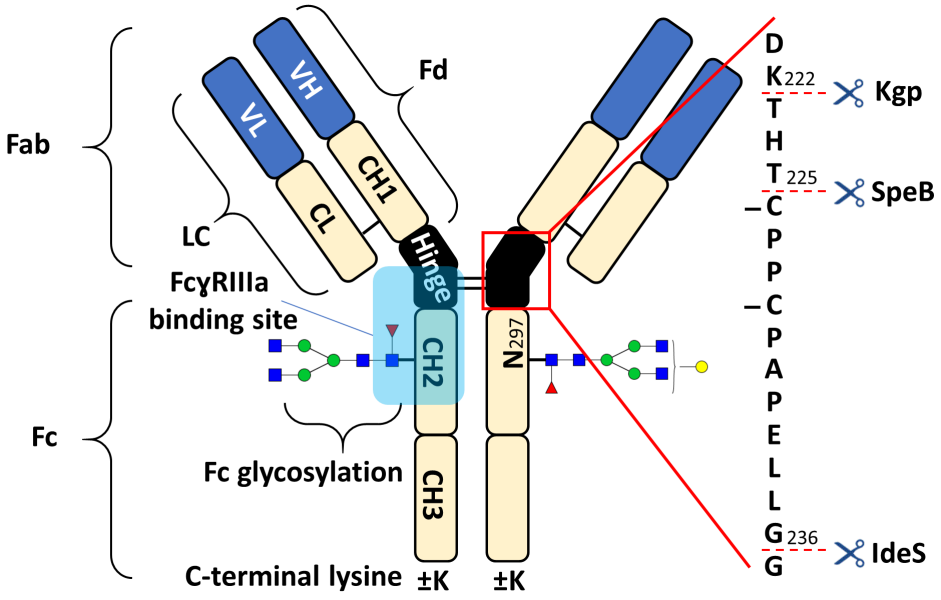

Figure 6.1: Schematic overview of human IgG1 with a zoom into the hinge region, indicating cleavage sites of IdeS, SpeB, and Kgp. The heavy chain (HC) contains three constant domains (CH1-CH3) and a variable domain (VH), whereas the light chain (LC) has only one constant domain (CL) and a variable domain (VL). The Fd consists of VH and CH1. LC and Fd together form the antigen binding fragment (Fab). About 15% of plasma IgG contain one or more additional *N*-glycosylation sites in the variable domains [202]. CH2 and CH3 of the two heavy chains build the fragment crystallizable (Fc). Within the Fc is a conserved glycosylation site at N297 in the CH2 [29]. Between the CH1 and CH2 of the heavy chains is a mostly flexible hinge region and the four chains are covalently connected via disulfide bridges. Amino acids are presented with single letter code and numbered according to the Kabat system [26].

Cetuximab is an approved mAb with additional Fab glycosylation and ADCC is described as one mechanisms of action [203, 204]. Each heavy chain (HC) contains an *N*-glycosylation site at the Fab (N88) and at the Fc (N299) resulting in a high number of possible glycoforms. The proteoform heterogeneity of cetuximab is further increased by C-terminal lysine variants of the HC [61]. Hence, glycoform assignment of the heavily glycosylated cetuximab by intact mass analysis is hindered by a high degree of ambiguities [61, 205]. Middle-up approaches are highly advantageous alternatives for obtaining information about individual subunit (e.g., Fc, Fc/2, Fab) modifications, especially for complex formats [39, 73, 206]. Bacterial enzymes are important tools for

middle-up approaches, since they cleave IgG specifically within the hinge region. Robust and simple workflows for the middle-up analysis of (therapeutic) mAbs are established [74, 75, 172, 205, 207]. IdeS, SpeB, and Kgp are frequently used commercial IgG hinge-specific proteases (cleavage sites and products are indicated in **Figure 6.1** and **Supplementary Figure 1**, respectively). As opposed to papain, for example, additional cleavages outside of the hinge region are not reported under standard conditions. Their characteristics were recently summarized [74]. IdeS based middle-up MS analysis of cetuximab is commonly applied to unravel the (glycosylation) microheterogeneity [61, 205, 208].

This study combines our recently reported FcγRIIIa AC-MS with middle-up analysis. Therefore, we investigated how cleavages within the hinge region affect the FcγRIIIa binding properties of the obtained Fc. Three different commercial IgG hinge-specific proteases were tested, namely IdeS, SpeB, and Kgp [74]. We demonstrate comparability of middle-up and intact affinity assessment by FcγRIIIa AC-MS upon Kgp digestion. Furthermore, we applied this workflow to cetuximab and simultaneously assessed the FcγRIIIa affinity, characterized the Fc glycoform pairings and analyzed the Fab glycosylation.

6.2. Materials and Methods

6.2.1. Chemicals, Proteases and Antibodies

All chemicals in this study had at least analytical grade quality. Deionized water was obtained from a Purelab ultra (ELGA Labwater, Ede, The Netherlands). Preparation of mobile phase was performed with ammonium acetate solution (7.5 M, Sigma-Aldrich, Steinheim Germany) and glacial acetic acid (Honeywell, Seelze, Germany). IdeS (FabRICATOR®), SpeB (FabULOUS®), and Kgp (GingisKHAN®) proteases were purchased from Genovis (Lund, Sweden). A reference standard therapeutic mAb produced in CHO cells (referred to as mAb1) and the FcγRIIIa affinity column was obtained from Roche Diagnostics (Penzberg, Germany). An EMA-approved cetuximab (Erbitux®) was used in this study. Cetuximab is a chimeric IgG1, produced by SP2/0 murine myeloma cells, which binds to the epidermal growth factor receptor (EGFR).

6.2.2. Antibody Digestion (IdeS, SpeB, Kgp)

All IgG hinge-specific proteases were reconstituted in deionized water following the manufacturer's instructions (IdeS: 67 units/µL, SpeB 40 units/µL, Kgp: 10 units/µL). Buffers and reducing conditions were selected from the recommended options. mAbs were digested at a concentration of 5 mg/mL (1 unit of protease per 1 µg of mAb) and incubated for 1 h at 37°C. In case of Kgp, the samples were buffer exchanged prior to digestion (10 kDa molecular weight cut-off filter, Merck, Darmstadt, Germany) to digestion buffer. 100 mM Tris buffer (pH 8) was used with mild reducing conditions (2

mM cysteine) for Kgp and reducing conditions (1 mM DTT) for SpeB, respectively. IdeS digestion was performed under non-reducing conditions with 100 mM ammonium bicarbonate (pH 7). After digestion, samples were buffer exchanged to a final concentration of 5 mg/mL in 50 mM ammonium acetate pH 5 (10 kDa molecular weight cut-off filter).

6.2.3. FcγRIIIa Affinity Chromatography-Mass Spectrometry

The FcγRIIIa AC-MS system was previously described in detail [198]. In short, a biocompatible Thermo Ultimate3000 instrument coupled to a 15 T solariX XR FT-ICR mass spectrometer (Bruker Daltonics, Bremen, Germany) was used. The column was operated at 25°C and a flow rate of 500 µL/min was applied. Prior to MS detection, the flow rate was reduced to 30 µL/min via flow-splitting. Mobile phase A was 50 mM ammonium acetate pH 5 and mobile phase B 50 mM ammonium acetate pH 3. Upon injection (50-100 µg sample), the column was washed for 10.5 min (5 column volumes) with 100% mobile phase A and then a linear gradient to 42 min to 100% mobile phase B (15 column volumes) was applied. For electrospray ionization (ESI), the capillary voltage was set to 4,000 V, the nebulizer gas to 0.8 bar, the dry gas flow to 3 L/min dry gas and the source temperature to 200°C.

6.2.4. Data Analysis

Average masses were calculated using the web-based Protein Tool (`https://www.protpi.ch`) based on the protein sequences and expected modifications. For mAb1 and cetuximab, C-terminal lysine clipping and 16 disulfide bridges were set as modifications. N-terminal pyroglutamic acid was additionally added as modification for cetuximab. Visualization and processing of mass spectra was performed in DataAnalysis 5.0 (Bruker Daltonics). Extracted ion chromatograms (EICs) were generated based on the theoretical *m/z* (± 0.2 Th) for all observed charge states. For deconvolution, the Maximum Entropy tool was used (deconvolution range indicated in table headings, data point spacing = 1, instrument resolving power = 3,000). All described Fc glycans can be found in **Supplementary Table 1** which provides information about composition and structure.

6.3. Results and Discussion

6.3.1. IgG Protease Evaluation

The FcγRIIIa AC-MS retention profiles of hinge cleaved mAb1, obtained by either IdeS, SpeB, or Kgp, and of intact mAb1 were compared (**Figure 6.2**). Although digestion sites of the three proteases are in close proximity in the hinge region (**Figure 6.1**), vastly different retention profiles were observed for the differently cleaved Fc. Kgp generated Fc was found to exhibit a remarkably comparable retention profile to the intact mAb1. IdeS digested mAb1 did not show retention on the FcγRIIIa column and the expected cleavage products, including the Fc, were detected in the injection peak (**Supplementary Figure 2**). Under native conditions, Fc fragments consisting of paired polypeptide chains were observed rather than single Fc/2 chains which is attributable to non-covalent interactions of the Fc polypeptides [205]. The lack of retention can be explained by the removal of amino acids that form an essential part of the FcγRIIIa binding motif [180]. In particular, L234 and L235 are crucial amino acids. The mutation of these amino acids to alanines (LALA mutant) is known to eliminate FcγRIIIa binding and thus ADCC [197, 209].

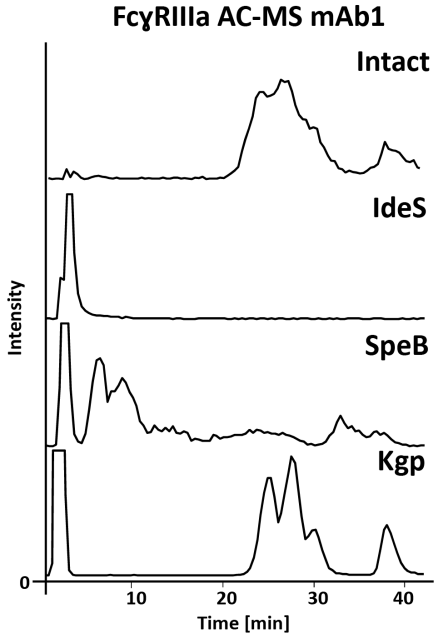

Figure 6.2: FcγRIIIa AC-MS base peak chromatograms of intact and IdeS-, SpeB-, and Kgp-digested mAb1. For intact and Kgp-digested mAb1, two distinct groups of peaks were detected, namely fucosylated forms with lower affinity (RT 20-35 min) and afucosylated glycoforms with higher affinity (RT higher than 35 min); see **Figure 3**.

In contrast to IdeS, the protease SpeB does not remove these key amino acids from the Fc. The Fab was observed in the injection peak while the Fc was retained on the FcγRIIIa column (**Supplementary Figure 3**). However, in contrast to Kgp, the Fc retention profile upon SpeB cleavage was vastly different from that of the intact mAb. SpeB derived Fc spread over the entire chromatogram and most of the Fc eluted already before the pH gradient started. Two differences from Kgp derived Fc might provide an explanation. Firstly, the removed amino acids (THT) might lead to an impaired conformational stability of the Fc obtained by SpeB. Furthermore, substitution of H, at this position, was shown to influence FcγRIIIa binding and ADCC [210]. Finally, partial reduction, which is likely to occur under the applied reducing conditions of 1 mM DTT [74], is discussed to influence binding to Fc receptors and ADCC [211]. Under milder reducing conditions SpeB does not show sufficient activity (data not shown). Besides the impaired retention profile, an increased heterogeneity is a disadvantage for MS analysis, when comparing SpeB to Kgp. Heterogeneity is caused by additional cleavages between H224 and T225 or T223 and H224 (**Supplementary Figure 3**). These products showed a similar impaired binding behavior. In contrast to SpeB, Kgp protease retains T223, H224, and T225 and works under mild reducing conditions (2 mM cysteine) which prevents reduction of the hinge interchain disulfide bonds [75]. Both, the additional amino acids and intact disulfide bonds, might be responsible for improved binding of Kgp derived Fc over SpeB derived Fc. Interestingly, the Fab was removed in a previous study using papain to exclude that the Fab influences the FcγRIIIa binding [181]. In line with our observations for Kgp in FcγRIIIa AC-MS, binding of papain generated Fc was comparable to intact IgG (glycovariants). Total binding strength as well as glycoform differences were preserved in the SPR analysis [181]. Papain digestion is performed under conditions (5-10 mM cysteine) which are only slightly more reductive than for Kgp (2 mM cysteine) [75, 181]. The preferred cleavage site of papain is between the Kgp and SpeB cleavage site (H224 and T225) [212, 213] and corresponds to the observed additional cleavage site of SpeB. Hence, the harsher incubation conditions are more likely to cause the differences between Kgp, or papain, and SpeB than the presence or absence of H224. However, by-products (due to cleavage outside the hinge region), insufficient yields and glycoform dependency make papain less favorable as an IgG middle-up protease [75, 214]. Moreover, papain might degrade the receptor material of the affinity column, since it is not an IgG-specific protease. Consequently, due to the preservation of the affinity separation and the specificity, Kgp was chosen for the middle-up FcγRIIIa AC-MS workflow. Interestingly, an IdeS middle-up approach is described for neonatal Fc receptor (FcRn) AC [44]. In this study, an impact of the Fab on the FcRn interaction was shown for several mAbs. The influence of different Fabs might also be relevant for other Fc receptor interactions. Thus, middle-up AC-MS also has high potential for investigating Fab-Fc structure-function relationships in a proteoform-resolved manner.

6.3.2. Comparability of Intact and Middle-Up FcγRIIIa Affinity Chromatography

FcγRIIIa AC, separates two distinct groups of peaks, representing fucosylated glycoforms (2x fucosylated glycoforms) with lower affinity and (partially) afucosylated glycoforms (2x and 1x afucosylated glycoforms) with higher affinity (**Figures 6.2, 6.3**) [198]. It has to be noticed that the column performance was slightly different for late eluting glycoforms compared to previous experiments on intact mAb1 [198]. Fc glycoforms are discussed in a nomenclature as listed in **Supplementary Table 1**. Middle-up FcγRIIIa AC-MS exhibited sharper peaks than its intact counterpart (**Figure 6.3**).

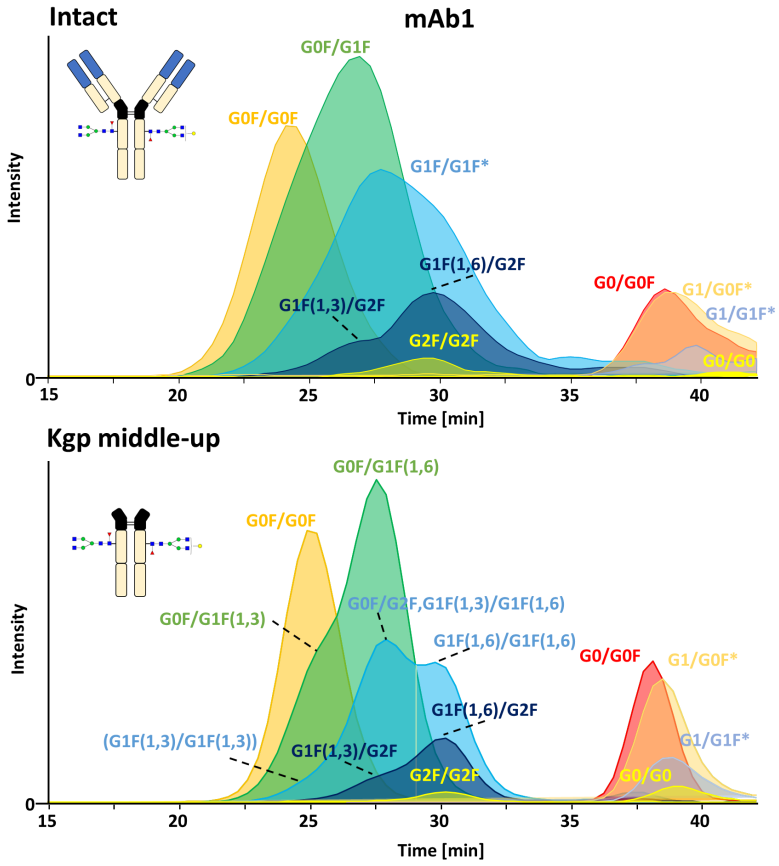

Figure 6.3: Extracted ion chromatograms of major glycoforms assigned for FcγRIIIa AC-MS of mAb1 on intact and Kgp middle-up level. *indicates additional isomers listed in **Supplementary Table 2**.

In general, decreasing the size of proteins in chromatography, in this case from 150 to 50 kDa, increases the diffusion coefficient [215]. This improves the mass transfer kinetics of the protein [216, 217] Sharper peaks and a similar retention resulted in better separation efficiency for the middle-up FcγRIIIa AC-MS. Mainly, several partially separated species in the EICs of **Figure 6.3** indicate separation of glycoforms with terminal galactose present on the 1,3-arm or the 1,6-arm of G1F. Differential FcγRIIIa affinity of G1F(1,3) and G1F(1,6) glycoforms has recently been reported [G2F = G1F(1,6) > G1F(1,3) = G0F] [218]. Based on this, G0F/G1F(1,3) has a similar affinity as G0F/G0F. G0F/G1F(1,6) exhibits an increased FcγRIIIa affinity comparable to G0F/G2F. Similarly, the G1F/G1F (G0F/G2F) peak shows partial separation. The first peak was assigned to G0F/G2F and G1F(1,3)/G1F(1,6) while the second peak likely represents G1F(1,6)/G1F(1,6). In addition, a third species [G1F(1,3)/G1F(1,3)] might populate the front of the peak. However, in this case peak asymmetry might provide an alternative explanation. For G1F/G2F, an early eluting peak [G1F(1,3)/G2F] was observed for the middle-up as well as the intact analysis. The affinity of G1F(1,6)/G2F was similar to G2F/G2F. A comparable glycoform ranking for mAb1 glycoforms is achieved in middle-up and intact FcγRIIIa AC-MS analysis. The masses corresponded to the expected Fc fragments (**Supplementary Table 2**) and comparability of glycoform affinity ranking to the intact mAb1 was demonstrated by comparing retention time differences (**Figure 6.3** and **Supplementary Figure 4**).

6.3.3. FcγRIIIa Affinity Analysis of Cetuximab Proteoforms

FcγRIIIa AC-MS of intact cetuximab is illustrated in **Figure 6.4** by EICs using *m/z* values of the three most abundant fucosylated and afucosylated Fc glycoforms, assuming H7N4F1/H7N4F1 as Fab glycosylation. Fab glycans are presented and discussed at a compositional level to avoid confusion with Fc glycans. Mainly fucosylated complex and high mannose glycoforms with little terminal galactose-α-1,3 galactose (α-gal) are described for the Fc glycosylation of SP2/0 produced cetuximab. For the Fab glycosylation, highly heterogeneous complex glycans with a high amount of α-gal and *N*-glycolylneuraminic acid (S) are reported [29, 219]. Intact cetuximab analysis proposed G0F/G1F and M5/G1F as the main fucosylated and (partially) afucosylated glycoforms, respectively (**Figure 6.5**). Based on literature, G0F/G0F and M5/G0F should be the most abundant Fc fucosylated and (partially) afucosylated glycoforms, respectively [61, 205]. This was later confirmed by middle-up analysis. FcγRIIIa AC-MS of intact cetuximab reduced the MS spectral complexity by separating fucosylated from (partially) afucosylated Fc glycoforms (**Figure 6.5**) compared to previously reported intact analysis [61, 205]. However, one MS peak may still be assigned to various combinations of Fc and Fab glycoforms with the same mass or similar masses. The degree of overlapping glycoforms and the resulting assignment ambiguities were exemplified by permutating three high abundant Fab glycoforms (H6N4F1, H6N4F1S1, H7N4F1) with the three most abundant Fc glycoforms (G0F, G1F, M5) already resulting in 36 different combinations (**Figure 6.5** and **Supplementary Table 3**). This is excluding heavy chain positional isomers within one site, such as G0F-H6N4F1/G1F-H7N4F1 and

G0F-H7N4F1/G1F-H6N4F1.

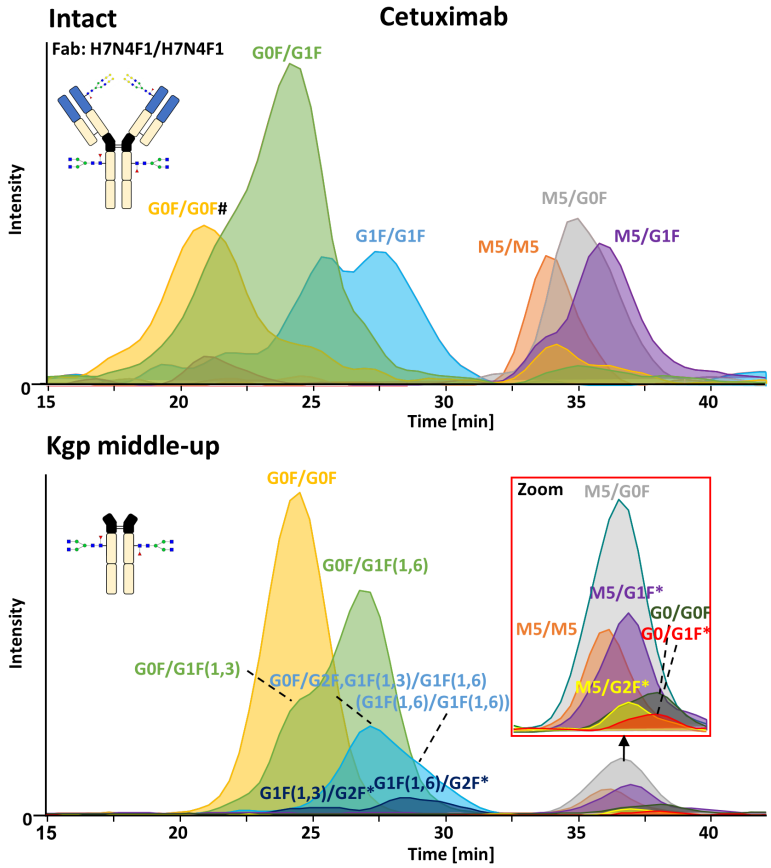

Figure 6.4: FcγRIIIa AC-MS of cetuximab on intact and Kgp middle-up level. Intact analysis was restricted to the main Fab glycoform (H7N4F1/H7N4F1) and the most abundant Fc glycoforms (G0F, G1F, M5). G0F/G0F# is marked exemplarily for isomeric EICs: G0F/G1F exhibits the same EIC if the Fab glycoform is H7N4F1/H6N4F1. For middle-up FcγRIIIa AC-MS analysis, EICs of all detected Fc glycoforms are presented. Data for C-terminal lysine variants (+K) are omitted for clarity. *indicates additional isomers listed in **Supplementary Table 4**.

The theoretical number of possible glycoforms is significantly higher when considering all possible Fc and Fab glycans. Thus, at the intact level, assessing FcγRIIIa affinity of cetuximab glycoforms by AC-MS based on EICs is prevented by the high number of isomeric and non-resolved proteoforms.

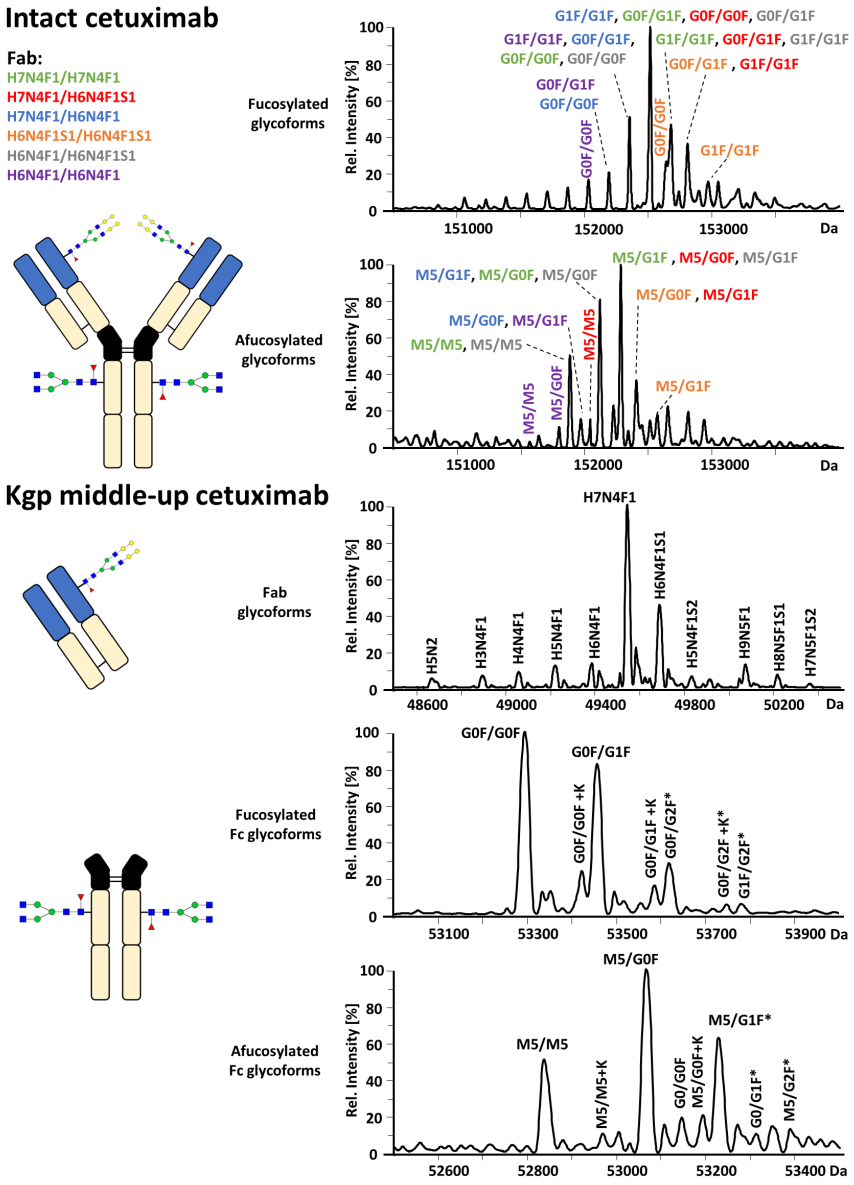

Figure 6.5: Assignment of cetuximab glycoforms on intact and Kgp middle-up level by FcγRIIIa AC-MS. Fc glycoform assignment of intact cetuximab was restricted to 6 different Fab glycoform pairings indicated by a color code. The broad peaks often contain multiple isomeric and non-resolved glycoforms (**Supplementary Table 3**). Thirty four out of thirty six possibilities are assigned to peaks. For Kgp middle-up FcγRIIIa AC-MS, the composition of Fab glycoforms were assigned as well as the Fc fragments (*indicates additional isomers). All observed deconvoluted masses for intact and middle-up analysis are listed in **Supplementary Tables 3-5**.

For example, Fc glycoforms G0F/G0F and G0F/G1F will be extracted as the same mass, when combined with Fab glycans of composition H7N4F1/H7N4F1 and H7N4F1/H6N4F1, respectively. The MS analysis of mAbs is generally affected by an increased signal heterogeneity derived from additional non-resolved proteoforms and adducts [220]. In particular for native MS of complex proteins, the applied deconvolution (algorithm, settings) can have an influence on the data evaluation with respect to resolving heterogeneous mass spectra [220]. Cetuximab glycoforms, containing H7N4F1/H6N4F1S1 and H6N4F1/H6N4F1S1 Fab glycoform pairings (**Figure 6.5** and **Supplementary Table 3**), were not resolved. A mass difference of 17 Da with H7N4F1/H7N4F1 and H6N4F1/H7N4F1 means the *m/z* difference for the most abundant charge state (28+) is 0.6 Th. This difference cannot be resolved under the applied conditions. Additionally, isomeric and non-resolved proteoforms lead to the distortion of the relative abundances as mentioned at the start of this paragraph. This becomes quite apparent when comparing the intact to the middle-up analysis in **Figures 6.4, 6.5**.

In contrast, middle-up FcγRIIIa AC-MS of cetuximab simplified MS spectra enough to confidently assign Fc glycoforms and lysine variants (**Figures 6.4, 6.5**). Nonetheless, the Fc glycoform pairing assignments also showed some degree of ambiguity (**Supplementary Table 4**, e.g., M5/G1F vs. M6/G0F). However, these ambiguities were minor compared to the intact mass analysis of cetuximab. Relative abundancies were in line with literature on Fc/2 glycoforms [61, 205]. The Fc glycan pairing of cetuximab was so far only briefly mentioned in a recent study, using IdeS digestion and direct infusion with native MS conditions [205]. We observed G0F/G0F as the main fucosylated glycoform. Additional galactosylation variants were observed as for mAb1. M5/G0F was determined as main (partially) afucosylated glycoform. Further high mannose glycoforms (M5/M5, M5/G1F, M5/G2F) were detected. Low amounts of G0/G0F, G0/G1F were also found. Relative FcγRIIIa affinity was comparable to mAb1 (**Figure 6.3**) and/or consistent with literature [198]. For example, a strong decrease and a mild increase in FcγRIIIa binding was observed for fucosylation and galactosylation, respectively [34, 181]. High mannose glycoforms exhibited a higher affinity than the fucosylated glycoforms but their affinity is decreased compared to the afucosylated complex-type glycoforms [192]. Interestingly, different pairings of high mannose glycoforms with fucosylated complex-type glycans (M5/G1F, M5/G2F) could be studied and were found with slightly higher FcγRIIIa affinity compared to M5/M5. M5/G1F and M5/G2F showed a slightly increased affinity over M5/G0F. This confirms and extends our previous findings, which were limited to the comparison of low abundant M5/M5 and M5/G0F [198]. Though it is difficult to compare affinity of the intact and the Fc, the latter seems to show a higher affinity (**Figure 6.4**). This is comparable to the negative influence of the Fab on FcRn binding reported for cetuximab [44].

Furthermore, C-terminal lysine variants (+K, **Figure 6.1**) could be studied in the middle-up analysis with respect to FcγRIIIa affinity. Investigations of incomplete lysine clipping with respect to different Fc glycoforms and FcγRIIIa affinity have not been described yet [221]. The C-terminal lysine appeared not to influence the Fc glycoform

retention strongly (**Supplementary Table 4**) which is not surprising as the receptor binds far away from the C-terminus (**Figure 6.1**). In contrast, C-terminal lysine may interfere with complement activation. However physiological relevance may anyhow be small as the lysine is enzymatically removed upon administration [222].

Moreover, middle-up FcγRIIIa AC-MS allowed a simultaneous determination of Fab glycosylation by evaluating the injection peak (**Figure 6.5** and **Supplementary Table 5**). Fab glycosylation might be relevant for antigen binding and pharmacokinetic behavior [223, 224]. In total, 11 Fab glycoforms were assigned. Our results are qualitatively and quantitatively in line with reported Fab glycosylation of SP2/0 produced cetuximab [61, 205]. H7N4F1 was determined to be the most abundant glycoform followed by H6N4F1S1, in line with a previous report [61] and in contrast to a recent study (H7N4F2) [205]. The two differently reported Fab glycoforms vary only by 1 Da. However, based on in-depth structural studies on cetuximab glycosylation reporting a high amount of S and only minor abundancies of antenna fucosylation, Fab glycoforms are more likely to contain H6N4F1S1 than additional H7N4F2 [225].

By applying middle-up FcγRIIIa AC-MS to cetuximab, 10 Fc glycoforms could be assigned, of which 5 Fc glycoforms were also found with the C-terminal lysine (amounting to 21 possible isomers, **Supplementary Table 4**). The combination of the 21 assigned Fc glycoform pairings (**Supplementary Table 4**) with the 11 assigned Fab glycoforms (**Supplementary Table 5**), belonging to 66 Fab glycoform pairings in theory, would result in 1,386 proteoforms for the intact cetuximab. This underlines the limitations of intact FcγRIIIa AC-MS for Fab glycosylated IgGs and highlights the degree of simplification obtained by the middle-up workflow.

6.4. CONCLUSION

IgG Fc, generated by three established IgG middle-up proteases (IdeS, SpeB, Kgp), demonstrated differences in retention behavior in FcγRIIIa AC-MS. Kgp derived Fc showed a remarkably similar retention profile and glycoform ranking as the intact mAb. Advantages of the Kgp middle-up FcγRIIIa AC-MS workflow were demonstrated in the application to the Fab glycosylated therapeutic mAb, cetuximab. The middle-up workflow provided significantly reduced MS complexity compared to the intact level. Consequently, it revealed important information about cetuximab Fc glycoforms by enabling their confident assignment and quantitation while retaining the Fc pairing and FcγRIIIa AC retention. Simultaneously, Fab glycosylation could be determined.

In the future, middle-up AC-MS might also be a tool to investigate the interplay of the Fab and the Fc in structure-function studies of IgG-Fc receptor interactions. Since the presented workflow with Kgp is limited to human IgG1, further development of hinge-specific antibody proteases with broader subclass coverage would be highly desired. AC-MS workflows may thus be further expanded toward clinical applications and polyclonal therapeutic samples (e.g., intravenous IgGs). Moreover, the complexity

of antibody-derived therapeutics is growing. Hence, the interest in middle-up approaches might further evolve in order to characterize new generations of therapeutic proteins.

SUPPORTING INFORMATION

Supporting information is available free of charge via https://www.frontiersin.org/articles/10.3389/fchem.2019.00698/full#supplementary-material.

ACKNOWLEDGEMENTS

This research was supported by the European Union (Analytics for Biologics project, Grant No. 765502), as well as by the NWO (Vernieuwingsimpuls Veni Project No. 722.016.008).

7

FC GAMMA RECEPTOR IIIB BINDING OF INDIVIDUAL ANTIBODY PROTEOFORMS RESOLVED BY AFFINITY CHROMATOGRAPHY–MASS SPECTROMETRY

Steffen Lippold, [1] Alexander Knaupp,[2] Arnoud H. de Ru, [1] Rayman T. N. Tjokrodirijo, [1] Peter A. van Veelen, [1] Erwin van Puijenbroek, [3] Steven W. de Taeye, [4] Dietmar Reusch, [5] Gestur Vidarsson, [4] Manfred Wuhrer[1] Tilman Schlothauer[2,6] and David Falck[1]

[1]Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands
[2]Pharma Research and Early Development, Roche Innovation Center Munich, Germany
[3]Pharma Research and Early Development, Roche Innovation Center Zurich, Switzerland
[4]Department of Experimental Immunohematology, Sanquin Research and Landsteiner Laboratory, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
[5]Pharma Technical Development, Roche Innovation Center Munich, Germany
[6]Biological Technologies, Genentech Inc. , South San Francisco, California

Reprinted and adapted from mAbs, in press
DOI: 10.1080/19420862.2021.1982847

The crystallizable fragment (Fc) of immunoglobulin G (IgG) activates key immunological responses by interacting with Fc gamma receptors (FcγR). FcγRIIIb contributes to neutrophil activation and is involved in antibody-dependent cellular cytotoxicity (ADCC) and antibody-dependent cellular phagocytosis (ADCP). These processes present important mechanisms-of-actions of therapeutic antibodies. The very low affinity of IgG towards FcγRIIIb ($K_D \approx 10\ \mu M$) is a technical challenge for interaction studies. Additionally, the interaction is strongly dependent on IgG glycosylation, a major contributor to proteoform heterogeneity. We developed an affinity chromatography – mass spectrometry (AC-MS) assay for analyzing IgG-FcγRIIIb interactions in a proteoform-resolved manner. This proved to be well suited to study low-affinity interactions. The applicability and selectivity of the method were demonstrated on a panel of nine different IgG monoclonal antibodies (mAbs), including no-affinity, low-affinity and high-affinity Fc-engineered or glycoengineered mAbs. Thereby, we could reproduce reported affinity rankings of different IgG glycosylation features and IgG subclasses. Additional post-translational modifications (IgG1 Met252 oxidation, IgG3 hinge-region O-glycosylation) showed no effect on FcγRIIIb binding. Interestingly, we observed indications of an effect of the variable domain sequence on the Fc-binding that deserves further attention. Our new AC-MS method is a powerful tool for expanding knowledge on structure-function relationships of the IgG-FcγRIIIb interaction. Hence, this assay may substantially improve the efficiency of assessing critical quality attributes of therapeutic mAbs with respect to an important aspect of neutrophil activation.

7.1. INTRODUCTION

IMMUNOGLOBULIN (Ig) G mediates key immunological responses by interacting with Fc gamma receptors (FcγR) [32]. FcγRIII is found mainly on macrophages, natural killer cells and neutrophils where it initiates various immune responses upon binding to opsonized IgG. The IgG-FcγRIII interaction is strongly glycosylation-dependent [35]. This is attributed to unique glycan-glycan and glycan-protein interactions between the receptor and the crystallizable fragment (Fc) of IgGs [31]. Fcs with an afucosylated *N*-glycan show drastically increased FcγRIII affinity [31, 226]. Increased affinity to activating FcγRs, such as FcγRIII, results in increased cytotoxicity [227]. Knowledge about the FcγR-IgG interaction enabled the rational design of anti-cancer monoclonal antibodies (mAb), glycoengineered for increased cytotoxicity [29].

FcγRIIIb is a particularly interesting receptor because it is uniquely expressed in humans. Neutrophils, the most abundant phagocytes in the circulation, show high levels of FcγRIIIb expression; in fact, the highest of any FcγR on any cell type [228]. Neutrophils exert antibody-dependent cellular cytotoxicity (ADCC), as well as antibody-dependent cellular phagocytosis (ADCP) [36]. The neutrophil activation via FcγRIIIb is considered an important mechanism of action of mAbs and is affected by the glycosylation, a critical quality attribute, of therapeutic mAbs [12, 36, 48, 229]. The extracellular domain of FcγRIIIb is highly homologous to FcγRIIIa (> 97% sequence homology [230]). However, FcγRIIIb is the only FcγR lacking a transmembrane and cytosolic signaling domain and is, instead, anchored by glycosylphosphatidylinositol. Further, the IgG1 affinity of FcγRIIIb ($K_D \approx 10$ μM) is up to ten-fold lower than for FcγRIIIa ($K_D \approx 1$ μM), which was attributed to a single amino acid difference [229, 230]. Of note, the K_D values are highly dependent on the mAb glycoform. IgG subclass specificity of FcγRIIIb interactions has been reported, with a higher affinity for IgG3 than IgG1 and no binding for IgG2 and IgG4 [227, 231]. Three polymorphic variants of FcγRIIIb, namely NA1, NA2 and SH, are known [227]. The two most common variants, NA1 and NA2, differ in four amino acids, leading to four (Asn38, Asn74, Asn162, Asn169) or six (Asn38, Asn45, Asn64 Asn74, Asn162, Asn169) glycosylation sites for NA1 and NA2, respectively. For IgG1 binding, only minor differences were observed for NA1 and NA2 [232].

In vitro measurements of monovalent affinity have been acknowledged as important metrics in mAb optimization [229]. Of note, FcγRIIIb affinity differences are difficult to measure with common techniques due to the low affinity and a high assay variability for FcγRIIIb affinity assessments [25, 233]. Various studies have assessed the effect of mAb glycosylation on FcγRIIIb affinity [35, 229, 232–234]. Besides fucosylation, galactosylation or bisecting N-acetylglucosamine (bisection) were found to modulate the interaction as well, but to a smaller extent. The naturally occurring heterogeneity of mAb glycosylation (i.e., glycan features and glycan pairing) is a major challenge for linking affinity differences to specific glyco- or proteoforms in most assays.

We recently developed an affinity chromatography – mass spectrometry (AC-MS) platform for a glycoform-resolved FcγRIIIa binding assessment of mAbs [198, 235]. In contrast to AC-UV [181], individual glycoforms within a complex mixture could be analyzed in a single run by AC-MS, omitting the need for glycoengineering. Furthermore, AC-MS enabled unpreceded insights into typically low-abundant glycoform pairings, which have not been addressed by previous studies. The molecular resolution obtained by AC-MS is an outstanding advantage over established physicochemical techniques such as surface plasmon resonance (SPR). In addition, retention time shifts in AC were previously linked to differences in ADCC activities [34]. This suggests that relevant approximations of the *in vivo* situation can be made by AC. Of note, physicochemical affinity assessment methods have high robustness and resolution. However, they do not fully reflect the *in vivo* complexity of biological interactions. Cell-based assays allow a better representation of the intricacies associated with immune complexes, at the expense of speed, robustness and resolution [179]. Therefore, these approaches are highly complementary and usually go hand in hand in drug discovery efforts.

This study reports, for the first time, the use of an FcγRIIIb affinity column for the binding assessment of mAbs. We developed the AC conditions for compatibility with MS and showed the glycoform-resolved affinity profiles of several classical and glycoengineered mAbs with high amounts of bisected and hybrid-type glycans. Furthermore, we demonstrated the selectivity of the FcγRIIIb AC-MS towards different IgG subclasses.

7.2. Results

7.2.1. Recombinant Fc gamma receptor glycosylation

A human embryonic kidney (HEK) cell-produced FcγRIIIb (NA1) was used in this study. The NA1 variant of FcγRIIIb has four glycosylation sites (Asn38, Asn74, Asn162, Asn169). At least one peptide moiety for each glycosylation site was verified by MS/MS analysis (**Figure S1 – S4**). A comprehensive site-specific analysis revealed 43 (Asn38), 56 (Asn74), 81 (Asn162) and 25 (Asn169) compositions at the individual sites, amounting to 95 different glycan compositions in total (**Figure S5, Table S1**). Asn162 glycosylation is visualized in more detail due to its functional importance (**Figure 7.1**). The relative abundance of aglycosylated Asn162 glycopeptide was negligible (0.1%). Furthermore, glycan compositions and structures of highly abundant glycopeptides from recombinant (HEK) and previously reported human neutrophil-derived FcγRIIIb were compared (**Figure 7.1**) [236].

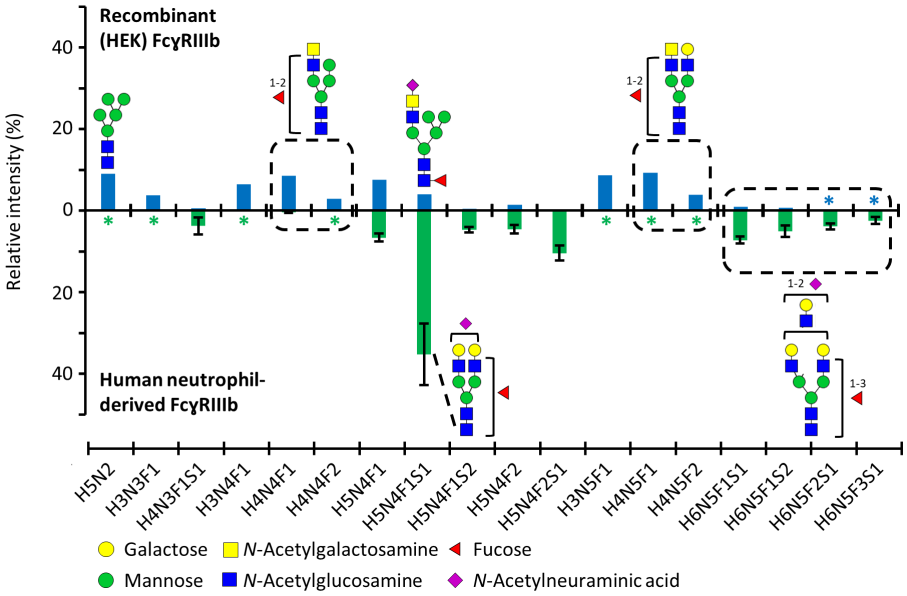

Figure 7.1: Asn162 glycosylation profiles of FcγRIIIb, comparing recombinant (HEK) and human neutrophil-derived receptor (extracted from Wojcik et al. [236]). The data was normalized to the sum of all quantified compositions.For recombinant (HEK) FcγRIIIb, these are listed in **Table S1**. Glycan structures of selected compositions are suggested based on MS/MS experiments and previous findings on HEK cell-derived or human-derived neutrophil FcγRIII glycosylation [189, 236]. Asteriks (*) indicate that a glycopeptide was not detected.

7.2.2. FC GAMMA RECEPTOR IIIB AFFINITY CHROMATOGRAPHY – MASS SPECTROMETRY

As mAbs have significantly lower affinity for FcγRIIIb than for FcγRIIIa, the separation was optimized starting from the previously reported conditions [181, 198]. Two mAbs were used for gradient optimization, representing either low afucosylation/low-affinity glycoforms (mAb1) or glycoengineered high afucosylation/high-affinity glycoforms (mAb3), respectively. A glycoengineered mAb with high levels of afucosylation showed up to seven-fold increased FcγRIIIb affinity compared to a mAb with a classical CHO-cell glycosylation profile, considering the average of all glycoforms [36]. Therefore, the glycoengineered mAb used for method development (mAb3) was expected to show higher affinity glycoforms compared to mAb1. The fucosylated species showed no retention using reported FcγRIIIa AC-MS and AC-UV conditions (**Figure S6**). Sufficient binding of fucosylated species was achieved by lowering the ammonium acetate concentration to 30 mM (mobile phase A). Further, the pH was increased to 6.8 and the column temperature decreased to 20°C, facilitating retention of fucosylated species. A complete elution within the gradient was obtained by an acetic

acid concentration of 50 mM in mobile phase B, which resulted in a final pH above 4 (**Figure 7.2**). For both mAbs, doubly fucosylated species could be chromatographically separated from singly and doubly afucosylated species. The charge state distribution of native MS spectra ($[M + 22H]^{22+}$ to $[M + 29H]^{29+}$) did not change within the gradient. This indicates that the mAbs do not undergo large conformational changes under the applied separation conditions (**Figure S7**). An Fc-engineered version of mAb1 was used as a negative control to check for non-specific interactions. This mAb1 version contained the Pro329Gly, Leu234Ala and Leu235Ala mutations, which abolish Fc receptor binding [209]. None of the glycoforms showed any retention (**Figure S8**).

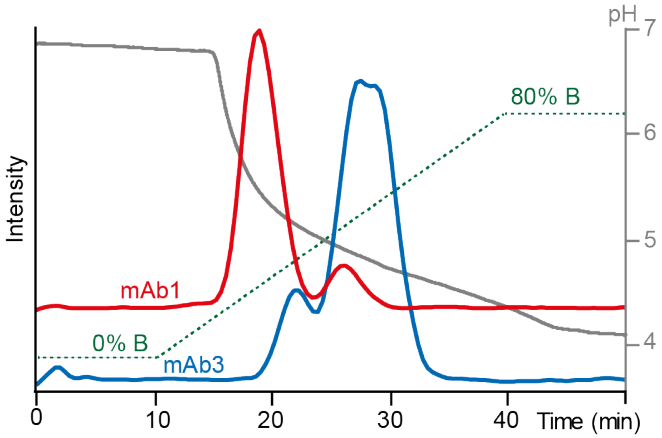

Figure 7.2: Optimized FcγRIIIb AC-MS gradient. 30 mM ammonium acetate pH 6.8 (mobile phase A) and 50 mM acetic acid pH 3.0 (mobile phase B) were used as mobile phases. Base peak chromatograms (BPCs) of mAb1 (red) and mAb3 (blue) are displayed. The gradient (green) and the pH (gray) are displayed as well.

7.2.3. Glycoform-resolved FcγRIIIb AC-MS of IgG1 mAbs

Four different IgG1 mAbs, produced in Chinese hamster ovary (CHO) cells, were analyzed in a glycoform-resolved manner by FcγRIIIb AC-MS (**Table S2, S3, Figure 7.3A-D**). Of note, the constant part of the heavy chains of mAb1, mAb3 and mAb4 is based on the G1m17 allotype (mAb1 = G1m17, mAb3/4 = G1m17,1), whereas mAb2 is based on the G1m3 allotype. The constant part of the light chains of all four mAbs is based on the Km3 allotype.

Doubly fucosylated (GxF/GxF), singly afucosylated (GxF/Gx) and doubly afucosylated (M5/M5, Gx/Gx) species were resolved by FcγRIIIb AC-MS. The assigned glycoforms of mAb1, affinity ranking and inter-day retention time variability showed high similarity to previously reported FcγRIIIa AC-MS results (**Table S3, Figure S9, Figure S10**) [198].

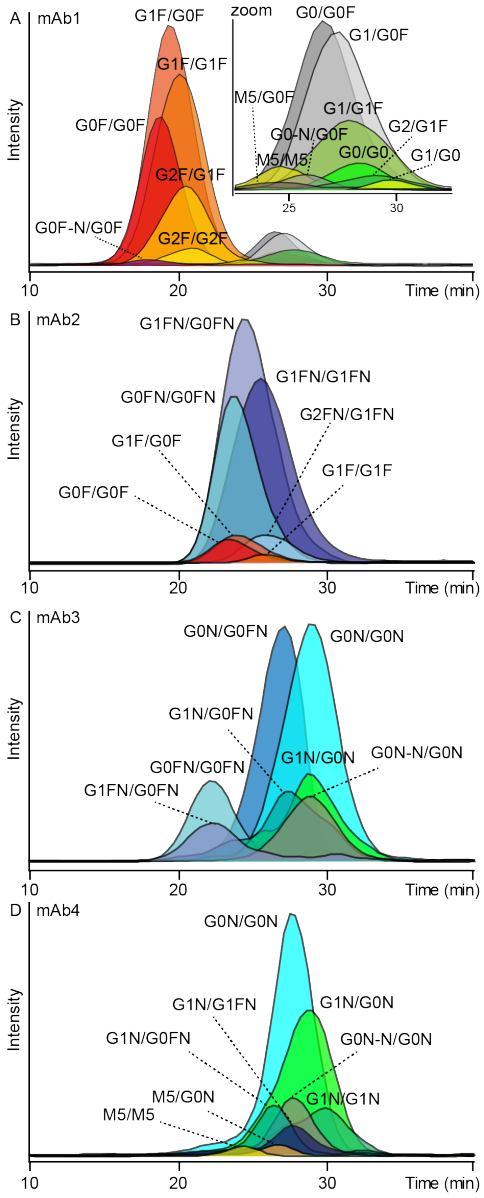

Figure 7.3: FcγRIIIb AC-MS analysis of four IgG1 mAbs. **A – D** represent mAb1 to mAb4. Extracted ion chromatograms of glycoforms are displayed (**Figure S9, S11 – 13, Table S3**). Zoom in A visualizes low-abundant afucosylated glycoforms of mAb1. Some assigned minor glycoforms (< 5% relative to the main peak) of mAb2 – mAb4 are not displayed for visibility reasons.

Next, the applicability of FcγRIIIb AC-MS to different glycoengineered mAbs (mAb2 – mAb4) was tested. Glycoforms of mAb2 were assigned mainly to fucosylated structures with varying levels of galactosylation and bisection (**Figure S11, Table S3**). Both galactosylation and bisection of mAb2 slightly increased retention in FcγRIIIb AC-MS. When comparing different IgG1 mAbs (**Figure 7.3**), the amino acid sequence apparently contributes to the retention. For example, the doubly fucosylated glycoforms (G0F/G0F, G1F/G0F, G1F/G1F) of mAb2 had increased retention compared to mAb1. In addition, the bisected fucosylated glycoforms of mAb2 (GxFN/GxFN) had increased retention compared to mAb3. Of note, the assigned glycoforms showed a preferential pairing of bisected glycoforms (GxFN/GxFN) or non-bisected glycoforms (GxF/GxF), whereas hemi-bisected glycoforms (GxFN/GxF) were not assigned. A minor amount of M5/M5 was detected and showed the highest retention time in mAb2 (**Figure S11, Table S3**).

mAbs glycoengineered via GlycoMab technology [237] (mAb3 and mAb4) showed predominantly bisected, afucosylated glycoforms (**Figure S12, S13, Table S3**). mAb3 showed bisected glycoforms with different fucosylation levels (2x, 1x, 0x). For mAb4, no doubly fucosylated glycoforms and overall higher afucosylation were observed. In addition, glycoforms consisting only of the conserved pentasaccharide core and a bisecting *N*-acetylglucosamine (G0N-N) were assigned, based on glycan data of the GlycoMab technology [237]. The influence of galactosylation was more pronounced for mAb4 (e.g., ΔG0N/G0N vs. G1N/G0N) compared to mAb3. Low abundant high mannose (M5/M5) species showed the lowest retention of afucosylated glycoforms for both mAbs.

In addition, a singly oxidized (Met252) variant of mAb4 was subjected to FcγRIIIb AC-MS (**Figure S14**). Enrichment was achieved by neonatal Fc receptor (FcRn) chromatography as described previously [45]. The glycoform masses of the oxidized version were in agreement with one additional oxidation. No notable influence of the oxidation on retention in FcγRIIIb AC-MS was observed (**Figure S14**).

7.2.4. IgG subclass specificity of FcγRIIIb visualized by AC-MS

We demonstrated that the reported IgG subclass specificity (IgG3 > IgG1, no retention of IgG2 and IgG4) of FcγRIIIb binding was reflected in the AC-MS approach. For this, a representative allotype of each of the four subclasses was analyzed. The different subclasses are referred to as IgG1 (G1m3; Km3), IgG2 (G2m(..); Km3), IgG3 (G3m(b*); Km3) and IgG4 (G4m(a); Km3). Similar glycosylation profiles and the same antigen-binding fragment (Fab) sequence (anti-trinitrophenol (TNP)) allowed direct comparison of glycoforms (**Figure 7.4A-D**). Mainly, fucosylated complex-type glycans with varying levels of galactosylation (GxF/GxF) were observed. As expected, IgG2 and IgG4 did not show retention on FcγRIIIb AC (**Figure 7.4B,D**).

The doubly fucosylated glycoforms of IgG1 (G1m3; Km3) showed comparable retention to the same glycoforms in mAb1 (G1m17; Km3). In addition to this main peak (16 – 25

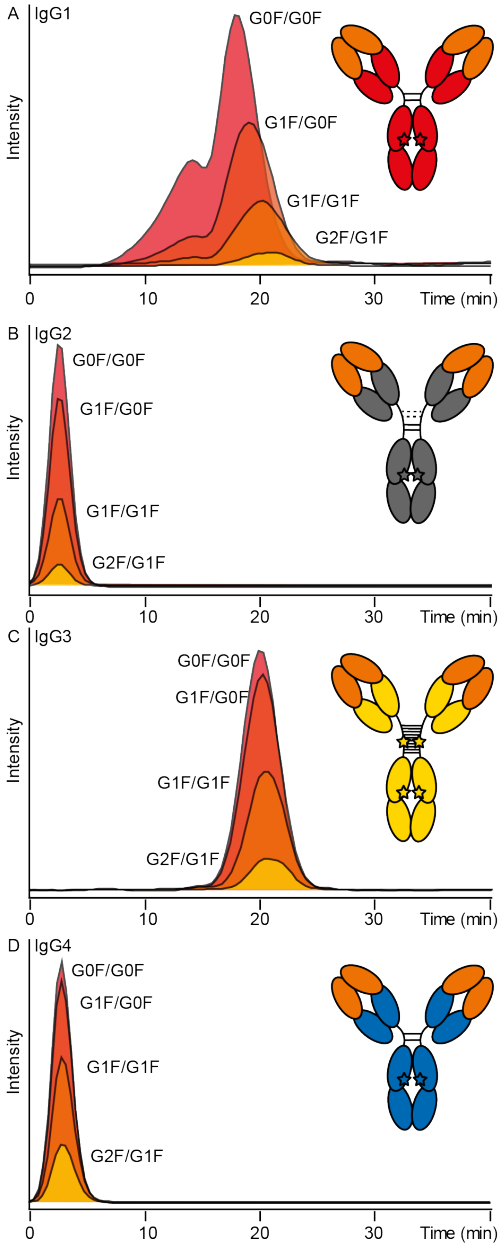

Figure 7.4: Anti-TNP IgG subclass comparison by FcγRIIIb AC-MS. Extracted ion chromatograms of the major glycoforms are displayed for IgG1-4 (in panels A-D, respectively).

min; peak 2), a partially separated peak with reduced retention (7 – 16 min; peak 1) was observed (**Figure 7.4A**). The two peaks did not show a difference in the intact mass profiles (**Figure S15**). Both peaks were fractionated in triplicates, analyzed by tryptic bottom-up analysis and checked for deamidation levels in the Fc (data not shown). However, the bottom-up data were inconclusive to link the partial separation to a potential Fc deamidation site.

IgG3 showed a slightly higher affinity compared to IgG1, but a less pronounced increase in retention time for galactosylation (**Figure 7.4A, C**). In addition to the main IgG3 glycoforms (GxF/GxF), glycoforms with *O*-glycosylation (+1x H1N1S2 or +2x H1N1S2) were assigned (**Figure S16**). However, no noticeable influence on the FcγRIIIb affinity was observed for the IgG3 hinge *O*-glycosylation (**Figure S17**).

7.3. DISCUSSION

7.3.1. RECEPTOR GLYCOSYLATION

FcγRIIIb glycosylation changes at Asn162 were reported to impact the interaction of FcγRIIIb with an IgG1 Fc fragment [234]. Therefore, it is imperative to report both the antibody and the receptor glycosylation, in order to facilitate the comparison of results between different interaction studies [227, 238]. The recombinant FcγRIIIb glycan profile of Asn162 showed differences compared to recently reported natural FcγRIIIb glycosylation (**Figure 7.1**). Recombinant vs. natural FcγRIIIb showed higher levels of oligomannose forms (9.9% vs. 1.7%), lower levels of sialylation (average number of sialic acids per glycan 0.2 vs. 0.9) and LacDiNAc structures were exclusively present on recombinant FcγRIIIb [238]. Furthermore, the amount of LacNAc units was less abundant in recombinant FcγRIIIb (**Figure 7.1**). FcγRIIIb from both sources showed antennary fucosylation (**Figure 7.1**, average number of fucoses per glycan 1.0 vs. 1.3). Interpreting the complex interplay between FcγRIIIb and antibody glycosylation remains challenging due to the scarcity of receptor glycosylation studies. Though we found glycosylation differences between recombinant and primary human FcγRs, there is as yet no indication that this would affect the affinity ranking of different antibody glycoforms. Nonetheless, the observed differences in oligomannose forms may well affect the absolute affinity of the antibodies [234, 239]. In conclusion, as we only report affinity rankings of IgG glycoforms, the recombinant FcγRIIIb is expected to provide state-of-the-art information. If, in the future, the use of recombinant FcγR with more human-like glycosylation profiles would be attainable and desirable, our method offers the possibility to quickly adapt to the potential changes in affinity.

7.3.2. SPECIFIC ADVANTAGES OF AFFINITY CHROMATOGRAPHY FOR LOW-AFFINITY INTERACTIONS

FcγRIIIb shows a very low affinity towards IgG1 ($K_D \approx 10$ μM) even when compared to other low-affinity FcγRs (FcγRIIIa, FcγRIIa) [25, 232]. For IgG1, a 10-fold lower affinity towards FcγRIIIb compared to FcγRIIIa was reported [23, 231]. The low affinity of the receptor-antibody interaction was reported to lead to assay-to-assay variation for FcγRIIIb binding [233]. A recent study applying SPR could not even determine binding affinities for different mAbs because the FcγRIIIb affinity was too weak [25]. In our study, AC conditions reported for FcγRIIIa showed insufficient retention for some glycoforms of the tested mAbs in FcγRIIIb AC. However, a generic FcγRIIIb AC gradient was developed with which both low and high-affinity mAb glycoforms could be analyzed. A pH gradient alone was not sufficient to achieve this broad coverage. Because we had observed a stronger susceptibility of the FcγRIIIb interaction to the buffer concentration, compared to FcγRIIIa, we successfully supplemented the pH gradient with a minor increase in buffer concentration. Of note, the pH gradient itself leads to a decrease in ionic strength. This likely counteracts the elution of antibodies, which is mitigated by increasing the buffer concentration with the gradient. This demonstrates one advantage of an AC-based assay for receptor antibody interaction studies. The chromatographic conditions offer much design space, while at the same time being tightly controlled. For example, different gradient parameters (ionic strength, pH) can be easily varied and the temperature is more easily controlled in a flow than in a batch setup. Consequently, AC-MS assays can be both highly adaptable and robust. A significant presence of non-specific binding, for example, through ionic interactions with the column material, was excluded using negative controls (mAb1 PGLALA variant, anti-TNP IgG2, anti-TNP IgG4). For different IgG1 mAbs and IgG subclasses, we could demonstrate that FcγRIIIb AC is a powerful and generic assay for assessing weak antibody-receptor interactions. It should be noted that the weak binding towards FcγRIIIb is expected to be less critical when studying multivalent interactions (avidity), i.e., of immune complexes, as opposed to monomeric mAb interactions (affinity). The affinity ranking of individual proteoforms is more insightful than solely K_D determination of proteoform mixtures by classical binding assays, such as SPR.

7.3.3. ADVANTAGES OF AC-MS FOR GLYCOFORM RESOLUTION

The online hyphenation of FcγRIIIb AC to native MS allows comprehensive proteoform assessment of heterogeneous mAbs and can resolve subtle affinity differences. We demonstrated recently the advantages of FcγRIIIa AC-MS [198]. In this and the previous study, mAb1 was comprehensively analyzed in a glycoform-resolved manner. The FcγRIIIb AC-MS glycoform affinity ranking and interday retention time stability of mAb1 showed high similarity compared to FcγRIIIa AC-MS (**Figure S10**). Of note, the data are not directly comparable, due to the need to adapt the chromatographic conditions to the lower FcγRIIIb affinity. The level of fucosylation had the biggest effect

on the FcγRIIIb affinity, which is attributed to a unique glycan-glycan interaction between antibody and receptor [31]. This is in line with other reports on FcγRIIIb-IgG interactions [35, 36, 229, 230, 233].

We observed differences in the effect of galactosylation in the studied mAbs. Of note, galactosylation has recently been found to affect the hexamerization potential of IgG1, which may possibly increase avidity contributions to FcγR interactions [240, 241]. We found no indication that other multivalent interactions, such as aggregate formation exist in our analytical setup. Interestingly, the effect of galactosylation in glycoengineered mAbs was more pronounced for mAb4 than for mAb3 (G0N/G0N vs G1N/G0N). This may be related to different linkages of galactosylation (presence of terminal galactose on the 1,3-arm or the 1,6-arm), which show a positive (1,6) or no effect (1,3) on FcγRIIIa binding [218, 235].

mAb2 showed the highest affinity of all doubly fucosylated and afucosylated (M5/M5) glycoforms (**Table S3**). A positive impact of the allotype can be excluded when comparing G0F/G0F of mAb2 with IgG1 (both G1m3; Km3). mAb2 was previously analyzed by FcγRIIIa AC-MS and the glycoforms showed increased RT as well when comparing to mAb1 [198]. However, the difference in FcγRIIIa affinity was not emphasized in the previous work. As similar FcγRIIIa binding profiles were reported, independent of the IgG1 allotype [25], the data hints to an influence of the variable domains on FcγRIII binding. This is in line with a recently demonstrated influence of the Fab on IgG1-FcγRIII receptor interaction [242, 243]. Although the Fab interaction sites were mainly attributed to the CH1 domain, amino acids within the variable domain were also found to contribute to the interaction [242]. However, further studies using orthogonal techniques are needed to validate the observed inter-mAb FcγRIIIb differences. If a strong impact of the variable Fab domain on FcγRIII would be confirmed, this may have very important implications for modulating mAb effector functions.

The glycoengineered mAbs featured several unstudied pairings, for example the combination of highly abundant G0N with oligomannose (G0N/M5) or bisected mono-antennary (G0N/G0N-N) glycoforms. Their investigation would require laborious glycoengineering, reducing, but not even fully eliminating, the heterogeneity of the mAb glycoform profiles. This highlights the advantage of FcγRIIIb AC-MS for affinity ranking of individual proteoforms. Here, we described the analysis of (therapeutic) mAbs produced in CHO or HEK cells, which are widely-used production systems [78]. Potential glycosylation differences to human IgG antibodies are extensively discussed in the literature [12]. The developed method is expected to apply to the majority of therapeutic mAbs. However, it should be noted that the proteoform resolution may be highly impaired for mAbs with additional complexity besides Fc glycosylation, e.g., Fab glycosylation. These may require additional method development as shown for the analysis of cetuximab by FcγRIIIa AC-MS [235]. Finally, it should be stressed that the developed assay targets qualitative differences between glycoforms. More precise insights may be obtained by comparing retention time

differences in AC-MS to quantitative data of highly homogeneous glycoengineered mAbs or fractions obtained by AC-MS.

7.3.4. AFFINITY IMPACT OF ADDITIONAL POST-TRANSLATIONAL MODIFICATIONS

Moreover, the glycoform-resolution obtained by FcγRIIIb AC-MS allowed study of the affinity impact of post-translational modifications (PTMs), other than Fc *N*-glycosylation. We concluded that oxidation at Met252 is not critical for FcγRIIIb binding (mAb4, **Figure S14**). In contrast, FcRn and FcγRIIa binding have been reported to be decreased in presence of Met252 oxidation [44, 45]. These differences with respect to Fc receptor binding are in line with the proximity to the different binding sites. In conclusion, we could demonstrate that AC-MS allowed the analysis of specifically enriched oxidized mAb variants for FcγRIIIb binding differences. Of note, PTMs cannot be located by intact mass analysis alone and the co-occurrence of multiple PTMs may hamper the data interpretation.

The partial separation of IgG1 glycoforms (**Figure 7.4A**) was hypothesized to be related to Asn deamidation in the Fc. Increased Asn325 deamidation levels were previously reported to reduce FcγRIIIa binding [43, 244]. However, additional experiments were inconclusive in supporting this interpretation. Partially reduced or scrambled disulfide bonds are another potential PTM, which is not resolved by intact mass analysis. This modification was also shown to affect FcγRIIIa binding [245]. Additional experiments would be needed to pinpoint the modification responsible for the additional separation of IgG1 proteoforms.

In contrast to the other IgG subclasses, some IgG3 allotypes additionally contain *O*-glycans. We could demonstrate the feasibility of measuring intact IgG3 with FcγRIIIb AC-MS. For the first time, we showed that the hinge *O*-glycosylation of IgG3 does not significantly influence the FcγRIIIb affinity. Moreover, the affinity of IgG3 was found to be slightly higher compared to IgG1, which is in line with recent findings [25]. IgG3-based biopharmaceuticals are not yet on the market, but the first clinical trials were recently started [23]. Despite its disadvantages, IgG3 has several benefits and may be reconsidered for biopharmaceuticals applications in the future. This makes a proteoform-resolved method for Fc receptor affinity assessments highly desirable [23].

The combination of FcγRIIIb AC and MS is very powerful to unravel complex glycosylated antibodies. It allows unpreceded molecular insights into the relative affinity ranking of unstudied glycoform pairings. We demonstrated the applicability to different (glycoengineered) mAbs. In addition, our data show the potential of AC-MS-based affinity assessments to simultaneously study the impact of different forms of post-translational modifications. All of this makes the method attractive for supporting the assessment of critical quality attributes.

7.3.5. IMPORTANCE OF UNDERSTANDING FCγRIIIB-MEDIATED EFFECTOR FUNCTIONS

The role of FcγRIIIb binding is less studied and understood compared to FcγRIIIa. Whereas the FcγRIIIa affinity of mAbs positively correlates with natural killer cell-mediated ADCC, recent studies demonstrated that glycoengineered/high-affinity mAbs increased ADCP, whereas ADCC was impaired [36, 246, 247]. In this context it is debated whether FcγRIIIb supports or hinders the FcγRIIa mediated activation of neutrophils [228, 246]. There is also evidence that FcγRIIIb by itself can lead to neutrophil activation [248]. Therefore, it is important to be aware of the relative expression levels and affinity of mAb proteoforms towards FcγRIIIb and FcγRIIa [249]. This could be crucial in understanding which neutrophil effector functions are activated by a therapeutic mAb. In addition, the use of high FcγRIII-affinity mAbs has been linked to safety concerns due to first infusion reactions, mediated by neutrophil FcγRIIIb [250]. Since FcγRIIIb is uniquely present in humans, these effects cannot be assessed by animal-based preclinical studies. In conclusion, an analytical platform for differentiating FcγRIIIb affinities of mAb proteoforms, as presented in this study, is of high relevance to better understand and predict the safety and efficacy of therapeutic mAbs.

7.4. MATERIALS AND METHODS

7.4.1. CHEMICALS

All chemicals were at least analytical grade and were purchased from Sigma-Aldrich (Steinheim, Germany), if not stated otherwise. A Purelab Ultra system (Veolia Water Technologies Netherlands B.V., Ede, Netherlands) system was used for deionized water. Mobile phases were prepared using an ammonium acetate solution (7.5 M) and glacial acetic acid (Fluka-Honeywell). First, a 1 M stock solution was prepared and further diluted to the target concentration. Proteases (GluC and chymotrypsin) were obtained from Worthington Biochemical Corp. (Lakewood, USA).

7.4.2. ANTIBODIES

Nine different monoclonal antibodies (mAbs) were used in this study. mAb1, mAb3 and mAb4, as well as a Pro329Gly, Leu234Ala and Leu235Ala mutant (PGLALA) of mAb1, were provided by Roche (Penzberg, Germany and Zurich, Switzerland). mAb2 (anti-RhD, reported as +B [35]) and anti-TNP subclasses (reported as IgG1*03, IgG2*01, IgG3*01, IgG4*03 [25]) were provided by Sanquin (Amsterdam, The Netherlands).

7.4.3. BOTTOM-UP LIQUID CHROMATOGRAPHY-MS/MS ANALYSIS

FcγRIIIb was subjected to in-gel digestion using GluC and chymotrypsin as reported previously [236]. Digested FcγRIIIb was subjected to bottom-up analysis by reversed-phase liquid chromatography (LC)-MS/MS. The separation of (glyco)peptides was performed on an Easy nLC 1200 system (Thermo Fisher Scientific) using a precolumn (15 mm × 100 µm; Reprosil-Pur C18-AQ 3 µm, Dr. Maisch, Ammerbuch, Germany) and an analytical nanoLC column (150 mm × 75 µm; Reprosil-Pur C18-AQ 3 µm). Mobile phase A was 0.1% formic acid in MQ. A gradient from 10% – 40% mobile phase B (0.1% formic acid/ 80% acetonitrile) was used for elution of (glyco-)peptides. The LC was hyphenated to an Orbitrap Exploris 480 mass spectrometer. MS1 scans were acquired in an *m/z* range of 400 – 3,500. The MS1 resolution was set to 120,000 (FcγRIIIb). Data-dependent higher-energy C-trap dissociation (HCD) was used for MS/MS fragmentation. An isolation window of 1.2 Th and a resolution of 30,000 was applied. The charge states 2 – 7 (FcγRIIIb) were included for fragmentation. For FcγRIIIb analysis, HCD with normalized collision energy (NCE) of 30% was performed. In addition, triggered MS/MS (HexNAc loss (204.087)) was used applying stepped NCE of 20%, 30% and 50% combined to one spectrum in an *m/z* range 110 – 3,500.

7.4.4. FCγRIIIB GLYCOPROTEOMIC DATA ANALYSIS

The obtained (glyco-)peptide cleavage products were verified by automated MS/MS identification using Byonic (v. 3.7.13 Protein Metrics). Next, glycopeptide compositions for each glycosylation site were analyzed based on mass accuracies and retention time differences (MS1 information) by GlycopeptideGraphMS [89]. A glycan list covering all glycans from the different glycosylation sites was generated, then applied to each glycopeptide portion, integrated and manually checked based on retention time, mass accuracy (< 10 ppm) and isotopic pattern quality (idotp > 0.85) in Skyline [123]. Relative abundances were calculated based on total area normalization for each glycosylation site.

7.4.5. FCγRIIIB COLUMN PREPARATION

The FcγRIIIb affinity column was prepared as reported for FcγRIIIa AC [34, 198]. In short, human FcγRIIIb_NA1 was produced in HEK cells as a construct with C-terminal AviTag, IgA protease cleavage site (no cleavage was performed) and IgG Fc part (PGLALA mutant) as reported [251]. The material was then biotinylated and 3 mg receptor were immobilized on streptavidin sepharose beads and packed in a Tricorn column housing (5 mm x 50 mm, GE Healthcare). The column volume was 1 mL.

7.4.6. FcγRIIIb AFFINITY LIQUID CHROMATOGRAPHY – MASS SPECTROMETRY

FcγRIIIb affinity chromatography was performed using a biocompatible Thermo Ultimate3000 instrument (Thermo Fisher Scientific). For method development, UV detection at 280 nm and online pH monitoring (PCM-3000) was used. The system was operated at 20°C and with a flow rate of 0.5 mL/min. Mobile phase A (30 mM ammonium acetate, pH 6.8) and mobile phase B (50 mM acetic acid, pH 3.0) was used. All samples were buffer exchanged to mobile phase A (30 kDa molecular weight cut-off filter, Merck, Darmstadt, Germany). 50 µg of the sample were injected for each run. Prior to injection, the column was conditioned for 30 min with mobile phase A. Upon injection, a washing step of 10 min mobile phase A was applied. Then a linear gradient of 30 min to 80% mobile phase B was used for elution. Next, an additional washing step with 80% mobile phase B for 10 min was used before returning to the starting conditions. The flow was split, diverting approximately 30 µL/min to the mass spectrometer. Online native MS detection was performed using a 15T solariX FT-ICR mass spectrometer (Bruker Daltonics, Bremen, Germany). MS settings were as described previously [198]. Acquired spectra were manually inspected and visualized using DataAnalysis 5.0 (Bruker Daltonics). Deconvolution of mass spectra was performed using the Maximum Entropy tool with an instrument resolving power of 3,000. The deconvolution mass range was set from 145,000 to 155,000 Da. Proteoform assignment was performed manually on deconvoluted mass spectra using a mass tolerance of 50 ppm. The web-based Protein Tool (`https://protpi.ch`) was used for calculation of average masses considering the mAb sequence, C-terminal lysine clipping, glycosylation, disulfide bonds and N-terminal pyro-glutamine (if applicable). Extracted ion chromatograms were generated for charge states 22+ to 29+ in a window of ± 0.4 Th.

SUPPORTING INFORMATION

The supplementary material is available free of charge via `https://www.tandfonline.com/doi/full/10.1080/19420862.2021.1982847`.

ACKNOWLEDGEMENTS

This research was supported by the Roche/Genentech Global Strategy Team (GST). David Falck received funding from the Dutch Research Council (NWO) in the framework of the ENW PPP Fund for the top sectors (project Proteoform-resolved pharmacokinetics of biopharmaceuticals, no. 019.012).

7

8

Discussion and Perspectives

THE safety and efficacy of therapeutic proteins are affected by glycosylation. Hence, the analysis and control of glycosylation have high relevance at all stages of biopharmaceutical drug development. However, glycosylation is highly heterogeneous and the main contributor to proteoform heterogeneity. A therapeutic protein may be glycosylated at multiple sites and exhibit a high variety of (isomeric) glycan structures. The inherent complexity is a major challenge for the in-depth and routine characterization of glycoproteins. Glycoform profiling methods are limited by the data analysis efficiency at the bottom-up level and a high degree of proteoform assignment ambiguities at the intact level. Thus, the integration of multilevel mass spectrometry (MS)-based strategies is pivotal for comprehensive glycosylation analysis. Linking specific glycoforms to functional differences cannot be achieved by commonly applied binding assays due to the proteoform complexity. Affinity separation in combination with MS may offer a way to form such links, relying on its sensitivity and capability to resolve individual proteoforms.

8.1. TOWARDS EFFICIENT GLYCOFORM PROFILING OF THERAPEUTIC PROTEINS

8.1.1. BOTTOM-UP GLYCOPEPTIDE ANALYSIS

The major bottleneck in bottom-up approaches for the analysis of isolated glycoproteins is data analysis [85, 252, 253]. Efficient identification and quantification of glycopeptides are challenged primarily by the degree and dynamic range of glycosylation microheterogeneity. Furthermore, unspecific cleavage products, isomeric/isobaric glycopeptides, and the co-occurrence of other PTMs hamper data analysis [85, 89]. Manual identification, aiming for comprehensive and accurate assessment of glycopeptide data, is tedious, error-prone, and time-consuming. Recent developments in bioinformatic tools enormously facilitate the automated analysis of glycopeptides [86, 87, 254, 255]. Besides a vast amount of MS/MS-centered tools, characteristic elution patterns of glycopeptides in reversed-phase liquid chromatography (RP-LC) have been exploited for developing more simplified and extensive MS1-centered data analysis [89, 90, 256]. The suitability of RP-LC MS for the integration of glycopeptide and peptide analysis makes it highly useful for multiple-attribute monitoring methods in biopharmaceutical applications as demonstrated in **Chapter 3** for the parallel assessment of glycosylation, Asn deamidation, and Met oxidation [122, 257].

The integration of software tools is key for efficient glycoproteomic data analysis. Platforms, such as OpenMS, offer already highly versatile and advanced data analysis pipelines for proteomics approaches [108]. However, the integration of glycopeptide-centered software tools is still emerging, even though many glycopeptide tools exist [252, 253, 255]. In **Chapter 2**, the integration of automated MS1-based glycopeptide identification (GlycopeptideGraphMS [89]) and data curation (LaCyTools [90]) was aided by a python script. This approach exemplifies a successful integration of independent glycopeptide-dedicated software tools towards more automated and

reliable data analysis. Glycosylation-centered databases, such as GlyConnect, are of high relevance to support glycoprotein characterization workflows [94]. For example, the existing knowledge on specific glycosylation of proteins or possible biosynthetic pathways can be easily integrated to interpret glycopeptide compositions [258]. Glycopeptide software tools ideally provide user interfaces, which facilitate the integration of other tools or databases. Moreover, visual interfaces, such as present in GlycopeptideGraphMS or GlyConnect, should be more widely used to simplify grasping the complexity of glycopeptide data. In a biopharmaceutical setting, it is most important that programs are well-maintained, easy to use, and centralized. The accurate assessment of PTMs such as glycosylation in an efficient manner is critical to accelerate the analytical characterization of therapeutic proteins. Commercial software tools (e.g. ProteinMetrics) for biopharmaceutical applications have largely emerged in recent years due to the demand for more comprehensive characterization and increasing complexity of biopharmaceuticals [85, 205, 259, 260]. In the future, the extension of commercial software packages by tools for simplified glycopeptide analysis workflows as discussed above is of high interest.

Although trypsin is the gold standard protease for (glyco)proteomic approaches, it may lead to glycopeptides covering multiple glycosylation sites owing to lacking cleavage sites [115]. This hampers the site-specific glycosylation analysis and is commonly addressed by the use of alternative proteases with different cleavage specificities. However, other proteases are usually less specific compared to trypsin [68]. Cysteine aminoethylation is an analytical approach to introduce additional tryptic cleavage sites by converting cysteines into pseudo-lysines and a new biopharmaceutical application was demonstrated in **Chapter 3** [119]. This approach has been described for other glycoproteomic and biopharmaceutical applications [118, 120, 121]. However, aminoethylation is not very often reported compared to more established alkylation strategies such as carbamidomethylation, despite its equal simplicity and cost. Cysteine aminoethylation should be more explored for the structural characterization of glycoproteins and biopharmaceuticals. For example, it should be further explored, whether aminoethylated and oxidized cysteines may be efficiently cleaved by Lys-N (N-terminal cleavage), which may prevent ambiguous glycopeptide assignments, which impaired IgA glycopeptide analysis in **Chapter 2**. For monoclonal antibodies (mAbs), cysteine-specific cleavage might be highly interesting for obtaining specific information at a level between bottom-up and middle-up analysis. Cysteine-specific cleavage has been demonstrated upon blocking of lysines, cysteine aminoethylation, and subsequent Lys-C digestion for several proteins [261]. This approach potentially increases the sequence coverage because it reduces the number of small, hydrophilic peptides that are often missing in bottom-up approaches. Additionally, the site-specificity of PTM localization compared to middle-up analysis is expected to be increased. The cysteine-specific cleavage of a standard IgG1 mAb (e.g. trastuzumab) would result in ten peptides of the heavy chain and five peptides of the light chain ranging from 2 kDa to 8 kDa. This mass range makes the peptide sequences highly accessible for fragmentation techniques and the number of only 15 peptides allows for straightforward data interpretation.

8.1.2. INTACT GLYCOPROTEIN ANALYSIS

Intact mass analysis of isolated therapeutic glycoproteins is a powerful approach for fast profiling purposes and for assessing glycoprotein similarities. In addition, intact mass analysis of glycoproteins is better suited for the analysis of highly sialylated glycoproteins compared to bottom-up methods as shown in **Chapter 4** [59, 262, 263]. It should be noted that the proteoform assignment is highly impaired for complex glycoproteins, such as erythropoietin (EPO). This is due to the overlapping and unresolved glycoforms as elaborated in **Chapter 4**. Hence, the integration of bottom-up and intact MS analysis data is emerging for the characterization of biopharmaceuticals due to the complementarity [59, 79, 263]. Even the charge profile of an intact mAb was recently reconstructed based on bottom-up data [260]. This demonstrates how valuable the integration of multilevel MS analyses is for a more comprehensive structural characterization of therapeutic proteins.

Matrix assisted laser desorption ionization (MALDI) in combination with time-of-flight (TOF) or Fourier-transform ion-cyclotron-resonance (FT-ICR) MS is being used for the analysis of intact proteins [168, 172]. However, the proteoform-resolution of more commonly used TOF-MS instruments is very low compared to MALDI FT-ICR MS instruments for complex glycoproteins such as EPO as exemplified in **Chapter 4**. The use of MALDI for intact mass analysis has several advantages for application in the biopharmaceutical industry. MALDI is highly automatable, fast, and, compared to electrospray ionization (ESI), less susceptible to sample carry-over. More importantly, the mass spectra processing is simplified since predominantly one charge state is present at high *m/z* (around *m/z* 15,000 for EPO, **Chapter 4**). In contrast, the more routinely used ESI results in multiple, potentially overlapping charge states at lower *m/z* ranges, even for native ESI [264]. Because of this, ESI-MS spectra are usually deconvoluted, which may introduce artefacts [205, 265]. One disadvantage of the presented MALDI FT-ICR MS method, compared to ESI-MS and MALDI TOF MS, is the limited mass range. Although CsI-cluster ion transmission up to *m/z* 35,000 was achieved, a significant decrease in sensitivity was observed over *m/z* 20,000. This limits the current method to proteins up to 40 kDa considering doubly charged ions are sufficiently present as observed for EPO in **Chapter 4**. It would be interesting to systematically test MALDI conditions, e.g. different matrices and additives, that lead to higher charge states or to investigate how the protein backbone or structure influences the observed charge state. It was recently demonstrated that higher charge states of up to 6+ could be observed for intact IgG by adjusting the MALDI sample preparation conditions[266]. This is still a drastically lower charge state compared to native ESI-MS (lowest charge state for native ESI-MS of mAbs is around 20+ [267]). Hence, MALDI methods may offer increased spatial *m/z* resolution for monitoring important proteoform information in a single charge state, making it an attractive tool for the analysis of complex glycoproteins.

The use of a MALDI FT-ICR MS instrument with intermediate source pressure allowed for increased sialic acid stability and enhanced ion declustering. Instrumental adjustments, such as ion source pressure and declustering capabilities, of more commonly available MALDI TOF MS instruments, should enable to improve the intact analysis of complex proteins, such as EPO. This would make the technology more widely accessible for biopharmaceutical laboratories. Furthermore, the advantages of MALDI MS for tolerating impurities (salts and detergents) should be assessed in comparison to direct-infusion ESI-MS for biopharmaceutical applications. It is expected that the increased tolerance makes MALDI a superior choice for analyzing samples from complex matrices such as cell-culture supernatant or formulation buffer [268, 269]. This makes the method more versatile for applications in the biopharmaceutical industry.

Higher proteoform assignment confidence is achieved by a separation dimension prior to MS, enabling more extensive proteoform profiling [270]. Several separation strategies, such as capillary electrophoresis or liquid chromatography hyphenated to MS, have been developed in recent years for biopharmaceuticals [270, 271]. Ion exchange LC-MS methods have been established for antibodies and are great to resolve proteoforms [70, 272]. Another recently developed AEX-MS method could highly decrease the proteoform assignment ambiguities for EPO, which are discussed in **Chapter 4** [71]. The developed FcγRIII AC-MS methods (**Chapter 5 – 7**) are sensitive to glycosylation features, particularly fucosylation, and greatly decrease proteoform assignment ambiguities for mAb glycoforms. For example, the overlap of galactosylation and fucosylation differences (Δ16 Da) is a common drawback of direct analysis approaches and impairs the distinction of G0F/G0F and G0F/G1 or G1F/G0. FcγRIII AC-MS resolves these proteoforms, which increases the confidence and number of assigned glycoforms. Hence, it represents a highly valuable tool for comprehensive mAb glycoform profiling at the intact level.

8.2. FUNCTIONAL AFFINITY CHROMATOGRAPHY – MASS SPECTROMETRY

8.2.1. THE POTENTIAL FOR SUPPORTING CRITICAL QUALITY ATTRIBUTE ASSESSMENT

The understanding of structure-function relationships is highly important for defining critical quality attributes (CQAs) of therapeutic proteins. This is an important step of biopharmaceutical drug development [48]. However, it is not trivial to link a single proteoform to functional changes, due to the inherent complexity of modifications such as glycosylation [179]. Glycoengineering is commonly applied to enable studying glycosylation features more specifically. This is highly laborious, time-consuming, and still results in heterogeneous glycoform mixtures [8, 34, 35]. AC-MS, yet neglected as a technology for biopharmaceutical purposes, has the potential to greatly support the identification of CQAs. Fc gamma receptor - antibody interactions, which mediate key immunological responses, were not previously studied by AC-MS. The developed AC-MS methods in this thesis (**Chapter 5 – 7**) tremendously enhance the understanding of individual proteoform structure-function relationships.

Fc receptor AC is considered a functional separation due to the role of mAb-receptor interactions in biological processes. Increased retention in FcγRIIIa AC has been linked to enhanced ADCC activation [34, 251]. The ranking of glycoforms, studied by AC-MS in this work, was very well in line with reported affinity differences [28]. However, the relative AC-MS retention time differences of glycoforms cannot be easily translated to their impact on binding affinities or function. For this purpose, enriched proteoforms obtained by preparative AC should be subjected to functional cellular assays. This might allow correlating the retention in AC to biological differences and would be extremely helpful for interpreting the data and concluding on the biological impact of specific proteoforms.

Native separation techniques such as IEX have already been proven to effectively link charge variants of mAbs to differences in binding affinities and/or cellular responses [43, 270, 273]. This is commonly achieved by fractionation of separated proteoforms and subjection to functional assays. Subsequently, bottom-up methods are used to identify and locate functionally critical modifications of enriched proteoforms. In contrast to physico-chemical separation techniques such as IEX, the retention of proteoforms in AC-MS is already linked to functional differences for specific receptor interactions [34, 251]. This increases greatly the functional proteoform selectivity of AC-MS and simplifies the identification of critical modifications. To better understand the suitability of AC-MS for studying common PTMs beyond glycosylation, stressed samples should be analyzed in the future. For evaluating potential degradation products, samples can be stored under specific stress conditions (e.g. 37°C at pH 8.5 for facilitating deamidation) to induce a variety of modified proteoforms [42]. For example, recent studies showed a decreased FcγRIIIa binding and ADCC activity of deamidated mAbs [43, 244]. It would be very interesting to see whether these deamidated

proteoforms are separated and what effect the fucosylation level has on the receptor interaction of deamidated proteoforms.

PTMs resulting in minor or no shifts in mass, such as deamidation or isomerization, are not resolved by intact mass spectrometry of mAbs. In addition, chromatographic separation of intact mAbs may also not be sufficient. Hence, bottom-up approaches are needed to confidently identify and localize PTMs. Fractionation and offline bottom-up workflows are time-consuming and often hampered as degradation products, e.g. deamidation, are artificially introduced during long sample preparation times [274]. Hence, it would be very advantageous to use AC in combination with an automated multidimensional LC approach to identify and localize unresolved modifications at the bottom-up level. A recent report demonstrated the feasibility of Protein A-AC (first dimension) in combination with online reduction (second dimension), tryptic digestion (third dimension), and RP LC-MS (fourth dimension) for mAb analysis [275]. The use of FcγRIII AC as the first dimension in this analytical setup has the potential to resolve more modifications and subtle changes. Additionally, this may also be useful to expand the applicability of AC to more complex formats, which may suffer from high proteoform assignment ambiguities in intact mass analysis. Subsequently, critical modifications can be selected for tailoring bottom-up-centered multiple-attribute monitoring methods for high-throughput or routine applications.

Besides the separation set-up, advanced MS acquisition strategies would be attractive to provide structural information in AC-MS top-down experiments [276]. Several MS/MS fragmentation techniques (e.g. ultraviolet photodissociation or electron-transfer dissociation) have been applied for mAb analysis achieving comprehensive sequence coverage and the localization of labile PTMs. The broad peaks of AC-MS allow for sufficient time to apply (multiple) fragmentation techniques. In particular, the analysis of Fc moieties (**Chapter 6**) is expected to result in high sequence coverage of hinge or CH2 modification sites, which are potentially critical for FcγRIIIa interactions.

For expanding the functional understanding of proteoform impact on relevant receptor interactions, the development of more AC columns is highly desirable. For example, mannose receptor or asialoglycoprotein interactions may assist the understanding of pharmacokinetics [277]. Moreover, it would be highly interesting to study Fab-related modifications using immobilized antigens for AC-MS. This is important for a functional understanding of proteoforms because critical modifications in the Fab (affecting antigen binding) are not separated in Fc receptor AC. For example, proteoforms with high Fc receptor binding, but low antigen-binding, may exhibit overall low efficacy. In the future, the analysis of mAbs using a repertoire of relevant AC columns could allow a more sound understanding of the functional implications of mAb proteoforms for CQA assessment.

8.2.2. TECHNICAL ADVANTAGES AND LIMITATIONS

Analytical AC columns with immobilized Fc receptors, such as FcRn, are commercialized and readily available. The specifications for these columns indicate high robustness, repeatability, and stability (> 100 injections). Even the readiness for clone selection in a high-throughput screening was recently demonstrated for an FcγRIIIa column [251]. Furthermore, analytical columns can be used for preparative purposes, i.e. enrichment of specific proteoforms for complementary functional or binding assays. This makes AC highly suitable for the functional assessment of mAbs in the biopharmaceutical industry. The reported AC methods are predominantly used in combination with UV detection [8, 34, 44, 181, 251]. This simplifies the analytical setup, but also drastically decreases the proteoform resolution.

The use of MS is needed for obtaining proteoform information. Detailed insights into individual proteoforms can be achieved by mass spectrometers optimized for native MS (high ion transmission and declustering). The Orbitrap with extended mass range (EMR) has been established in recent years within industrial laboratories and is expected to give a comparable data quality as obtained by the 15T FT-ICR MS instrument used in this work [278]. Moreover, MS methods are emerging in quality control laboratories, which demonstrates the readiness for MS-based methods beyond research and development [257, 279]. Intact mass analysis methods are currently expanding and are expected to become more established for routine applications.

The developed AC-MS methods focus on monomeric mAb-receptor interactions and high proteoform resolution. The technical robustness, sensitivity and precision of cell-free assays, such as AC-MS, comes at the cost of reducing the relevance to the *in vivo* situation [179]. In contrast, cellular assays are more representative of the *in vivo* situation but suffer from low robustness. Compared to AC-MS, they also lack proteoform resolution. In general, a positive correlation exists between the assay complexity and the physiological relevance. However, the analysis of underlying monomeric mAb - receptor interactions is of high relevance for understanding the molecular details of biological responses.

One of the most established cell-free binding assays for biopharmaceuticals is surface plasmon resonance spectroscopy (SPR) [179, 280, 281]. SPR methods are used for determining binding kinetics and are suitable for high-throughput applications [282, 283]. However, SPR by itself cannot study individual proteoforms. The integration AC-MS and SPR analyses could add highly relevant information for biomolecular interactions. For example, SPR may be used online with LC for determining binding kinetics in parallel to proteoform identification by MS as described recently [284]. An AC-SPR/MS approach could add very proteoform-specificbinding kinetics for understanding the impact of retention time differences in AC-MS. This approach may allow a more quantitative output of AC-MS methods for determining affinity differences. Moreover, the glycoforms resolved by FcγRIII AC may allow for probing of individual glycoforms with different receptors in the SPR dimension. This integrated

approach is expected to further enhance structure-function studies of mAb glycoforms for relevant receptor interactions.

Another very recently developed approach for studying mAb-receptor interactions in a proteoform-resolved manner is affinity capillary electrophoresis-MS (ACE-MS) (Gstöttner et al. manuscript in preparation). This method assesses affinity differences by shifts in electrophoretic mobility. ACE-MS allows for determining mAb binding kinetics by analyzing a variety of receptor concentrations. Moreover, this technique is very versatile and accessible as the receptor does not need to be immobilized and packed in a column. In addition, the required amount of sample material is much lower (only a few ng per analysis) compared to AC-MS. However, the migration time in CE is, compared to retention time in LC, less robust. In addition, proteoforms (e.g. charge variants) may show already differential migration in the absence of a receptor, which complicates the data analysis. ACE-MS requires higher technical expertise to operate the setup. In addition, CE-MS applications are generally less user-friendly [271, 285]. In contrast, the operation of an LC column can be easily performed routinely, which makes AC-MS more suitable for methods transfers. For use in biopharmaceutical research, ACE-MS could guide the selection of AC columns to study and monitor critical receptor interactions. Overall, AC(-MS) approaches are expected to have broader applicability and transferability within all stages of biopharmaceutical drug development.

ADDENDUM

REFERENCES

[1] G. Walsh, *Biopharmaceutical benchmarks 2018,* Nat. Biotechnol. **36**, 1136 (2018).

[2] G. W. H. Organization, *Who model list of essential medicines - 21st list, 2019,* (2019).

[3] G. Walsh, *Therapeutic insulins and their large-scale manufacture,* Appl. Microbiol. Biotechnol. **67**, 151 (2005).

[4] J. W. Eschbach, J. C. Egrie, M. R. Downing, J. K. Browne, and J. W. Adamson, *Correction of the anemia of end-stage renal disease with recombinant human erythropoietin,* N Engl J Med **316**, 73 (1987).

[5] T. M. Pierpont, C. B. Limper, and K. L. Richards, *Past, present, and future of rituximab—the world's first oncology monoclonal antibody therapy,* Front Oncol **8** (2018).

[6] H. Kaplon and J. M. Reichert, *Antibodies to watch in 2021,* mAbs **13**, 1860476 (2021).

[7] A. Mullard, *Fda approves 100th monoclonal antibody product,* **20**, 491 (2021).

[8] R. Wada, M. Matsui, and N. Kawasaki, *Influence of n-glycosylation on effector functions and thermal stability of glycoengineered igg1 monoclonal antibody with homogeneous glycoforms,* mAbs **11**, 350 (2019).

[9] L. Liu, *Antibody glycosylation and its impact on the pharmacokinetics and pharmacodynamics of monoclonal antibodies and fc-fusion proteins,* J. Pharm. Sci. **104**, 1866 (2015).

[10] C. E. Halstenson, M. Macres, S. A. Katz, J. R. Schnieders, M. Watanabe, J. T. Sobota, and P. A. Abraham, *Comparative pharmacokinetics and pharmacodynamics of epoetin alfa and epoetin beta,* Clin Pharmacol Ther **50**, 702 (1991).

[11] S. Hua, M. J. Oh, S. Ozcan, Y. S. Seo, R. Grimm, and H. J. An, *Technologies for glycomic characterization of biopharmaceutical erythropoietins,* Trends Analyt Chem **68**, 18 (2015).

[12] D. Reusch and M. L. Tejada, *Fc glycans of therapeutic antibodies as critical quality attributes,* Glycobiology **25**, 1325 (2015).

[13] K. Stavenhagen, H. Hinneburg, M. Thaysen-Andersen, L. Hartmann, D. Varón Silva, J. Fuchser, S. Kaspar, E. Rapp, P. H. Seeberger, and D. Kolarich, *Quantitative mapping of glycoprotein micro-heterogeneity and macro-heterogeneity: an evaluation of mass spectrometry signal strengths using synthetic peptides and glycopeptides,* J Mass Spectrom **48**, 627 (2013).

[14] P. Stanley, N. Taniguchi, and M. Aebi, *N-glycans,* in *Essentials of Glycobiology*, edited by A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, and P. H. Seeberger (Cold Spring Harbor Laboratory Press Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved., Cold Spring Harbor (NY), 2015) pp. 99–111.

[15] K. T. Schjoldager, Y. Narimatsu, H. J. Joshi, and H. Clausen, *Global view of human protein glycosylation pathways and functions,* Nat Rev Mol Cell Biol **21**, 729 (2020).

[16] I. Brockhausen, H. Schachter, and P. Stanley, *O-galnac glycans,* in *Essentials of Glycobiology*, edited by A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart, and M. E. Etzler (Cold Spring Harbor Laboratory Press Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California., Cold Spring Harbor (NY), 2009) Book section 9.

[17] T. Raju, J. B. Briggs, S. M. Borge, and A. J. S. Jones, *Species-specific variation in glycosylation of igg: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics,* Glycobiology **10**, 477 (2000).

[18] A. Rita Costa, M. Elisa Rodrigues, M. Henriques, J. Azeredo, and R. Oliveira, *Guidelines to cell engineering for monoclonal antibody production,* Eur J Pharm Biopharm **74**, 127 (2010).

[19] J. Dumont, D. Euwart, B. Mei, S. Estes, and R. Kshirsagar, *Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives,* Crit. Rev. Biotechnol. **36**, 1110 (2016).

[20] E. Maeda, S. Kita, M. Kinoshita, K. Urakami, T. Hayakawa, and K. Kakehi, *Analysis of nonhuman n-glycans as the minor constituents in recombinant monoclonal antibody pharmaceuticals,* Anal. Chem. **84**, 2373 (2012).

[21] L. K. Hmiel, K. A. Brorson, and M. T. Boyne, *Post-translational structural modifications of immunoglobulin g and their effect on biological activity,* Anal. Bioanal. Chem. **407**, 79 (2015).

[22] R. Jefferis, *Isotype and glycoform selection for antibody therapeutics,* Arch Biochem Biophys **526**, 159 (2012).

[23] T. H. Chu, E. F. Patz, and M. E. Ackerman, *Coming together at the hinges: Therapeutic prospects of igg3,* mAbs **13**, 1882028 (2021).

[24] R. Jefferis and M.-P. Lefranc, *Human immunoglobulin allotypes,* mAbs **1**, 332 (2009).

[25] S. W. de Taeye, A. E. H. Bentlage, M. M. Mebius, J. I. Meesters, S. Lissenberg-Thunnissen, D. Falck, T. Sénard, N. Salehi, M. Wuhrer, J. Schuurman, A. F. Labrijn, T. Rispens, and G. Vidarsson, *Fcγr binding and adcc activity of human igg allotypes,* Front Immunol **11** (2020).

[26] E. A. Kabat, T. Te Wu, H. M. Perry, C. Foeller, and K. S. Gottesman, *Sequences of proteins of immunological interest* (DIANE publishing, 1992).

[27] R. Jefferis, *Glycosylation of recombinant antibody therapeutics,* Biotechnol Prog **21**, 11 (2005).

[28] W. Shatz, S. Chung, B. Li, B. Marshall, M. Tejada, W. Phung, W. Sandoval, R. F. Kelley, and J. M. Scheer, *Knobs-into-holes antibody production in mammalian cell lines reveals that asymmetric afucosylation is sufficient for full antibody-dependent cellular cytotoxicity,* mAbs **5**, 872 (2013).

[29] R. Jefferis, *Glycosylation as a strategy to improve antibody-based therapeutics,* Nat Rev Drug Discov **8**, 226 (2009).

[30] C. Spahr, S. D. H. Shi, and H. S. Lu, *O-glycosylation of glycine-serine linkers in recombinant fc-fusion proteins,* mAbs **6**, 904 (2014).

[31] C. Ferrara, S. Grau, C. Jäger, P. Sondermann, P. Brünker, I. Waldhauer, M. Hennig, A. Ruf, A. C. Rufer, M. Stihle, P. Umaña, and J. Benz, *Unique carbohydrate–carbohydrate interactions are required for high affinity binding between fcγriii and antibodies lacking core fucose,* Proc Natl Acad Sci U S A **108**, 12669 (2011).

[32] F. Nimmerjahn and J. V. Ravetch, *Fcγ receptors as regulators of immune responses,* Nat Rev Immunol **8**, 34 (2008).

[33] R. L. Shields, J. Lai, R. Keck, L. Y. O'Connell, K. Hong, Y. G. Meng, S. H. A. Weikert, and L. G. Presta, *Lack of fucose on human igg1 n-linked oligosaccharide improves binding to human fcγriii and antibody-dependent cellular toxicity,* J Biol Chem **277**, 26733 (2002).

[34] M. Thomann, T. Schlothauer, T. Dashivets, S. Malik, C. Avenal, P. Bulau, P. Rüger, and D. Reusch, *In vitro glycoengineering of igg1 and its effect on fc receptor binding and adcc activity,* PLoS One **10**, e0134949 (2015).

[35] G. Dekkers, L. Treffers, R. Plomp, A. E. H. Bentlage, M. de Boer, C. A. M. Koeleman, S. N. Lissenberg-Thunnissen, R. Visser, M. Brouwer, J. Y. Mok, H. Matlung, T. K. van den Berg, W. J. E. van Esch, T. W. Kuijpers, D. Wouters, T. Rispens, M. Wuhrer, and G. Vidarsson, *Decoding the human immunoglobulin g-glycan repertoire reveals a spectrum of fc-receptor- and complement-mediated-effector activities,* Front Immunol **8** (2017).

[36] J. Golay, F. Da Roit, L. Bologna, C. Ferrara, J. H. Leusen, A. Rambaldi, C. Klein, and M. Introna, *Glycoengineered cd20 antibody obinutuzumab activates neutrophils and mediates phagocytosis through cd16b more efficiently than rituximab,* Blood **122**, 3482 (2013).

[37] A. Beck and J. M. Reichert, *Marketing approval of mogamulizumab: a triumph for glyco-engineering,* mAbs **4**, 419 (2012).

[38] D. Falck, M. Thomann, M. Lechmann, C. A. M. Koeleman, S. Malik, C. Jany, M. Wuhrer, and D. Reusch, *Glycoform-resolved pharmacokinetic studies in a rat model employing glycoengineered variants of a therapeutic monoclonal antibody,* mAbs **13**, 1865596 (2021).

[39] A. Beck, E. Wagner-Rousset, D. Ayoub, A. Van Dorsselaer, and S. Sanglier-Cianférani, *Characterization of therapeutic antibodies and related products,* Anal Chem **85**, 715 (2013).

[40] G. Walsh and R. Jefferis, *Post-translational modifications in the context of therapeutic proteins,* Nat Biotechnol **24**, 1241.

[41] P. Bults, R. Bischoff, H. Bakker, J. A. Gietema, and N. C. van de Merbel, *Lc-ms/ms-based monitoring of in vivo protein biotransformation: Quantitative determination of trastuzumab and its deamidation products in human plasma,* Anal Chem **88**, 1871 (2016).

[42] Q. Yan, M. Huang, M. J. Lewis, and P. Hu, *Structure based prediction of asparagine deamidation propensity in monoclonal antibodies,* mAbs **10**, 901 (2018).

[43] X. Lu, L. A. Machiesky, N. De Mel, Q. Du, W. Xu, M. Washabaugh, X.-R. Jiang, and J. Wang, *Characterization of igg1 fc deamidation at asparagine 325 and its impact on antibody-dependent cell-mediated cytotoxicity and fcγriiia binding,* Sci Rep **10**, 383 (2020).

[44] T. Schlothauer, P. Rueger, J. O. Stracke, H. Hertenberger, F. Fingas, L. Kling, T. Emrich, G. Drabner, S. Seeber, J. Auer, S. Koch, and A. Papadimitriou, *Analytical fcrn affinity chromatography for functional characterization of monoclonal antibodies,* mAbs **5**, 576 (2013).

[45] F. Cymer, M. Thomann, H. Wegele, C. Avenal, T. Schlothauer, D. Gygax, and H. Beck, *Oxidation of m252 but not m428 in hu-igg1 is responsible for decreased binding to and activation of hu-fcγriia (his131),* Biologicals **50**, 125 (2017).

[46] T. Čaval, W. Tian, Z. Yang, H. Clausen, and A. J. R. Heck, *Direct quality control of glycoengineered erythropoietin variants,* Nat Commun **9**, 3342.

[47] S. Kozlowski and P. Swann, *Current and future issues in the manufacturing and development of monoclonal antibodies,* Adv Drug Deliv Rev **58**, 707 (2006).

[48] N. Alt, T. Y. Zhang, P. Motchnik, R. Taticek, V. Quarmby, T. Schlothauer, H. Beck, T. Emrich, and R. J. Harris, *Determination of critical quality attributes for monoclonal antibodies using quality by design principles,* Biologicals **44**, 291 (2016).

[49] I. D. G. Campuzano, M. Nshanian, C. Spahr, C. Lantz, C. Netirojjanakul, H. Li, P. Wongkongkathep, J. J. Wolff, and J. A. Loo, *High mass analysis with a fourier transform ion cyclotron resonance mass spectrometer: From inorganic salt clusters to antibody conjugates and beyond,* J Am Soc Mass Spectrom **31**, 1155 (2020).

[50] R. J. Rose, E. Damoc, E. Denisov, A. Makarov, and A. J. R. Heck, *High-sensitivity orbitrap mass analysis of intact macromolecular assemblies,* Nat Methods **9**, 1084 (2012).

[51] M. E. Belov, E. Damoc, E. Denisov, P. D. Compton, S. Horning, A. A. Makarov, and N. L. Kelleher, *From protein complexes to subunit backbone fragments: A multi-stage approach to native mass spectrometry,* Anal Chem **85**, 11163 (2013).

[52] Z. Zhang, P. K. Chan, J. Richardson, and B. Shah, *An evaluation of instrument types for mass spectrometry-based multi-attribute analysis of biotherapeutics,* mAbs **12**, 1783062 (2020).

[53] M. Mann, *The ever expanding scope of electrospray mass spectrometry—a 30 year journey,* Nat Commun **10**, 3744 (2019).

[54] D. Reusch, M. Haberger, B. Maier, M. Maier, R. Kloseck, B. Zimmermann, M. Hook, Z. Szabo, S. Tep, J. Wegstein, N. Alt, P. Bulau, and M. Wuhrer, *Comparison of methods for the analysis of therapeutic immunoglobulin g fc-glycosylation profiles—part 1: Separation-based methods,* mAbs **7**, 167 (2015).

[55] V. Dotz, R. Haselberg, A. Shubhakar, R. P. Kozak, D. Falck, Y. Rombouts, D. Reusch, G. W. Somsen, D. L. Fernandes, and M. Wuhrer, *Mass spectrometry for glycosylation analysis of biopharmaceuticals,* Trends Analyt Chem **73**, 1 (2015).

[56] D. Rathore, A. Faustino, J. Schiel, E. Pang, M. Boyne, and S. Rogstad, *The role of mass spectrometry in the characterization of biologic protein products,* Expert Rev Proteom **15**, 431 (2018).

[57] F. Hillenkamp and M. Karas, *Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization,* in *Methods in Enzymology*, Vol. 193 (Academic Press, 1990) pp. 280–295.

[58] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse, *Electrospray ionization for mass spectrometry of large biomolecules,* Science **246**, 64 (1989).

[59] Y. Yang, F. Liu, V. Franc, L. A. Halim, H. Schellekens, and A. J. Heck, *Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity,* Nat Commun **7**, 13397 (2016).

[60] J. Giorgetti, A. Beck, E. Leize-Wagner, and Y.-N. François, *Combination of intact, middle-up and bottom-up levels to characterize 7 therapeutic monoclonal antibodies by capillary electrophoresis – mass spectrometry,* J Pharm Biomed Anal **182**, 113107 (2020).

[61] D. Ayoub, W. Jabs, A. Resemann, W. Evers, C. Evans, L. Main, C. Baessmann, E. Wagner-Rousset, D. Suckau, and A. Beck, *Correct primary structure assessment and extensive glyco-profiling of cetuximab by a combination of intact, middle-up, middle-down and bottom-up esi and maldi mass spectrometry techniques,* mAbs **5**, 699 (2013).

[62] M. Karas, D. Bachmann, U. Bahr, and F. Hillenkamp, *Matrix-assisted ultraviolet laser desorption of non-volatile compounds,* Int J Mass Spectrom Ion Processes **78**, 53 (1987).

[63] M. Karas and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons,* Anal Chem **60**, 2299 (1988).

[64] M. Karas and R. Krüger, *Ion formation in maldi: The cluster ionization mechanism,* Chem Rev **103**, 427 (2003).

[65] R. Knochenmuss and R. Zenobi, *Maldi ionization: The role of in-plume processes,* Chem Rev **103**, 441 (2003).

[66] F. Hillenkamp and Peter-Katalinic, *MALDI MS: a practical guide to instrumentation, methods and applications* (John Wiley & Sons, 2013).

[67] I. D. G. Campuzano and W. Sandoval, *Denaturing and native mass spectrometric analytics for biotherapeutic drug discovery research: Historical, current, and future personal perspectives,* J Am Soc Mass Spectrom (2021).

[68] P. Giansanti, L. Tsiatsiani, T. Y. Low, and A. J. Heck, *Six alternative proteases for mass spectrometry based proteomics beyond trypsin,* Nat Protoc **11**, 993 (2016).

[69] L. W. Dick, D. Mahon, D. Qiu, and K.-C. Cheng, *Peptide mapping of therapeutic monoclonal antibodies: Improvements for increased speed and fewer artifacts,* J Chromatogr B **877**, 230 (2009).

[70] M. Haberger, A.-K. Heidenreich, M. Hook, J. Fichtl, R. Lang, F. Cymer, M. Adibzadeh, F. Kuhne, H. Wegele, D. Reusch, L. Bonnington, and P. Bulau, *Multiattribute monitoring of antibody charge variants by cation-exchange chromatography coupled to native mass spectrometry,* J Am Soc Mass Spectrom (2021).

[71] G. van Schaick, C. Gstöttner, A. Büttner, D. Reusch, M. Wuhrer, and E. Domínguez-Vega, *Anion exchange chromatography – mass spectrometry for monitoring multiple quality attributes of erythropoietin biopharmaceuticals,* Anal Chim Acta **1143**, 166 (2021).

[72] R. Gahoual, A.-K. Heidenreich, G. W. Somsen, P. Bulau, D. Reusch, M. Wuhrer, and M. Haberger, *Detailed characterization of monoclonal antibody receptor interaction using affinity liquid chromatography hyphenated to native mass spectrometry,* Anal Chem **89**, 5404 (2017).

[73] J. Sjögren, F. Olsson, and A. Beck, *Rapid and improved characterization of therapeutic antibodies and antibody related products using ides digestion and subunit analysis,* Analyst **141**, 3114 (2016).

[74] J. Sjögren, L. Andersson, M. Mejàre, and F. Olsson, *Generating and purifying fab fragments from human and mouse igg using the bacterial enzymes ides, speb and kgp,* in *Bacterial Pathogenesis: Methods and Protocols*, edited by P. Nordenfelt and M. Collin (Springer New York, New York, NY, 2017) pp. 319–329.

[75] J. Moelleken, M. Endesfelder, C. Gassner, S. Lingke, S. Tomaschek, O. Tyshchuk, S. Lorenz, U. Reiff, and M. M. , *Gingiskhan™protease cleavage allows a high-throughput antibody to fab conversion enabling direct functional assessment during lead identification of human monoclonal and bispecific igg1 antibodies,* mAbs **9**, 1076 (2017).

[76] C. Fu, H. Zhao, Y. Wang, H. Cai, Y. Xiao, Y. Zeng, and H. Chen, *Tumor-associated antigens: Tn antigen, stn antigen, and t antigen,* HLA **88**, 275 (2016).

[77] Z. Szabo, J. R. Thayer, D. Reusch, Y. Agroskin, R. Viner, J. Rohrer, S. P. Patil, M. Krawitzky, A. Huhmer, N. Avdalovic, S. H. Khan, Y. Liu, and C. Pohl, *High performance anion exchange and hydrophilic interaction liquid chromatography approaches for comprehensive mass spectrometry-based characterization of the n-glycome of a recombinant human erythropoietin,* J. Proteome Res. **17**, 1559 (2018).

[78] N. de Haan, D. Falck, and M. Wuhrer, *Monitoring of immunoglobulin n- and o-glycosylation in health and disease,* Glycobiology **30**, 226 (2019).

[79] Y. Yang, V. Franc, and A. J. R. Heck, *Glycoproteomics: A balance between high-throughput and in-depth analysis,* Trends in Biotechnol **35**, 598 (2017).

[80] S. Pan, R. Chen, R. Aebersold, and T. A. Brentnall, *Mass spectrometry based glycoproteomics—from a proteomics perspective,* Mol Cell Proteomics **10** (2011).

[81] M. Thaysen-Andersen and N. H. Packer, *Advances in lc-ms/ms-based glycoproteomics: getting closer to system-wide site-specific mapping of the n- and o-glycoproteome,* Biochim. Biophys. Acta **1844**, 1437 (2014).

[82] O. Ozohanics, L. Turiák, A. Puerta, K. Vékey, and L. Drahos, *High-performance liquid chromatography coupled to mass spectrometry methodology for analyzing site-specific n-glycosylation patterns,* J Chromatogr A **1259**, 200 (2012).

[83] L. R. Ruhaak, G. Xu, Q. Li, E. Goonatilleke, and C. B. Lebrilla, *Mass spectrometry approaches to glycomic and glycoproteomic analyses,* Chem Rev **118**, 7886 (2018).

[84] J. Klein and J. Zaia, *Relative retention time estimation improves n-glycopeptide identifications by lc–ms/ms,* J. Proteome Res. (2020).

[85] L. Y. Lee, E. S. X. Moh, B. L. Parker, M. Bern, N. H. Packer, and M. Thaysen-Andersen, *Toward automated n-glycopeptide identification in glycoproteomics,* J. Proteome Res. **15**, 3904 (2016).

[86] J. L. Abrahams, G. Taherzadeh, G. Jarvas, A. Guttman, Y. Zhou, and M. P. Campbell, *Recent advances in glycoinformatic platforms for glycomics and glycoproteomics,* Curr. Opin. Struct. Biol. **62**, 56 (2020).

[87] H. Hu, K. Khatri, and J. Zaia, *Algorithms and design strategies towards automated glycoproteomics analysis,* Mass Spectrom Rev **36**, 475 (2017).

[88] M. Bern, Y. J. Kil, and C. Becker, *Byonic: Advanced peptide and protein identification software,* Curr Protoc Bioinformatics **40**, 13.20.1 (2012).

[89] M. S. Choo, C. Wan, P. M. Rudd, and T. Nguyen-Khuong, *Glycopeptidegraphms: Improved glycopeptide detection and identification by exploiting graph theoretical patterns in mass and retention time,* Anal. Chem. **91**, 7236 (2019).

[90] B. C. Jansen, D. Falck, N. de Haan, A. L. Hipgrave Ederveen, G. Razdorov, G. Lauc, and M. Wuhrer, *Lacytools: A targeted liquid chromatography–mass spectrometry data processing package for relative quantitation of glycopeptides,* J. Proteome Res. **15**, 2198 (2016).

[91] A. Momčilović, N. de Haan, A. L. Hipgrave Ederveen, A. Bondt, C. A. M. Koeleman, D. Falck, L. A. de Neef, W. E. Mesker, R. Tollenaar, A. de Ru, P. van Veelen, M. Wuhrer, and V. Dotz, *Simultaneous immunoglobulin a and g glycopeptide profiling for high-throughput applications,* Anal. Chem. **92**, 4518 (2020).

[92] R. Plomp, N. de Haan, A. Bondt, J. Murli, V. Dotz, and M. Wuhrer, *Comparative glycomics of immunoglobulin a and g from saliva and plasma reveals biomarker potential,* Front Immunol **9** (2018).

[93] M. Šimurina, N. de Haan, F. Vučković, N. A. Kennedy, J. Štambuk, D. Falck, I. Trbojević-Akmačić, F. Clerc, G. Razdorov, A. Khon, *et al.*, *Glycosylation of immunoglobulin g associates with clinical features of inflammatory bowel diseases,* Gastroenterology **154**, 1320 (2018).

[94] D. Alocci, J. Mariethoz, A. Gastaldello, E. Gasteiger, N. G. Karlsson, D. Kolarich, N. H. Packer, and F. Lisacek, *Glyconnect: Glycoproteomics goes visual, interactive, and analytical,* J. Proteome Res. **18**, 664 (2019).

[95] D. Falck, B. C. Jansen, N. de Haan, and M. Wuhrer, *High-throughput analysis of igg fc glycopeptides by lc-ms,* in *High-Throughput Glycomics and Glycoproteomics* (Springer, 2017) pp. 31–47.

[96] N. Deshpande, P. H. Jensen, N. H. Packer, and D. Kolarich, *Glycospectrumscan: fishing glycopeptides from ms spectra of protease digests of human colostrum siga,* J. Proteome Res. **9**, 1063 (2010).

[97] K. B. Chandler, N. Mehta, D. R. Leon, T. J. Suscovich, G. Alter, and C. E. Costello, *Multi-isotype glycoproteomic characterization of serum antibody heavy chains reveals isotype-and subclass-specific n-glycosylation profiles,* Mol Cell Proteomics **18**, 686 (2019).

[98] S. Zheng and A. A. Doucette, *Preventing n- and o-formylation of proteins when incubated in concentrated formic acid,* Proteomics **16**, 1059 (2016).

[99] J. Nilsson and G. Larson, *Sialic acid capture-and-release and lc-msn analysis of glycopeptides,* in *Mass Spectrometry of Glycoproteins: Methods and Protocols*, edited by J. J. Kohler and S. M. Patrie (Humana Press, Totowa, NJ, 2013) pp. 79–100.

[100] S. Joenväärä, I. Ritamo, H. Peltoniemi, and R. Renkonen, *N-glycoproteomics–an automated workflow approach,* Glycobiology **18**, 339 (2008).

[101] H. Steen and M. Mann, *Similarity between condensed phase and gas phase chemistry: Fragmentation of peptides containing oxidized cysteine residues and its implications for proteomics,* J Am Soc Mass Spectrom. **12**, 228 (2001).

[102] V. Gupta and K. S. Carroll, *Sulfenic acid chemistry, detection and cellular lifetime,* Biochim. Biophys. Acta **1840**, 847 (2014).

[103] Z. Darula and K. F. Medzihradszky, *Carbamidomethylation side reactions may lead to glycan misassignments in glycopeptide analysis,* Anal. Chem. **87**, 6297 (2015).

[104] D. G. Delafield and L. Li, *Recent advances in analytical approaches for glycan and glycopeptide quantitation,* Mol Cell Proteomics , mcp.R120.002095 (2020).

[105] Y. J. Kil, C. Becker, W. Sandoval, D. Goldberg, and M. Bern, *Preview: a program for surveying shotgun proteomics tandem mass spectrometry data,* Anal. Chem. **83**, 5259 (2011).

[106] K. R. Rebecchi, J. L. Wenke, E. P. Go, and H. Desaire, *Label-free quantitation: A new glycoproteomics approach,* J Am Soc Mass Spectrom **20**, 1048 (2009).

[107] M. Sturm, A. Bertsch, C. Gröpl, A. Hildebrandt, R. Hussong, E. Lange, N. Pfeifer, O. Schulz-Trieglaff, A. Zerck, K. Reinert, and O. Kohlbacher, *Openms – an open-source software framework for mass spectrometry,* BMC Bioinformatics **9**, 163 (2008).

[108] H. L. Röst, T. Sachsenberg, S. Aiche, C. Bielow, H. Weisser, F. Aicheler, S. Andreotti, H.-C. Ehrlich, P. Gutenbrunner, E. Kenar, X. Liang, S. Nahnsen, L. Nilse, J. Pfeuffer, G. Rosenberger, M. Rurik, U. Schmitt, J. Veit, M. Walzer, D. Wojnar, W. E. Wolski, O. Schilling, J. S. Choudhary, L. Malmström, R. Aebersold, K. Reinert, and O. Kohlbacher, *Openms: a flexible open-source software platform for mass spectrometry data analysis,* Nat. Methods **13**, 741 (2016).

[109] J. Byeon, Y.-R. Lim, H.-H. Kim, and J.-K. Suh, *Structural identification of a non-glycosylated variant at ser126 for o-glycosylation site from epo brp, human recombinant erythropoietin by lc/ms analysis,* Mol Cells **38**, 496 (2015).

[110] D. Falck, M. Haberger, R. Plomp, M. Hook, P. Bulau, M. Wuhrer, and D. Reusch, *Affinity purification of erythropoietin from cell culture supernatant combined with maldi-tof-ms analysis of erythropoietin n-glycosylation,* Sci Rep **7**, 5324 (2017).

[111] P. Lai, R. Everett, F.-F. Wang, T. Arakawa, and E. Goldwasser, *Structural characterization of human erythropoietin,* J. Biol. Chem. **261**, 3116 (1986).

[112] P. E. Groleau, P. Desharnais, L. CotÃ©, and C. Ayotte, *Low lc ms/ms detection of glycopeptides released from pmol levels of recombinant erythropoietin using nanoflow hplc chip electrospray ionization,* J Mass Spectrom **43**, 924 (2008).

[113] Y. Takegawa, H. Ito, T. Keira, K. Deguchi, H. Nakagawa, and S. Nishimura, *Profiling of n and o glycopeptides of erythropoietin by capillary zwitterionic type of hydrophilic interaction chromatography/electrospray ionization mass spectrometry,* J Sep Sci **31**, 1585 (2008).

[114] E. Giménez, R. Ramos-Hernan, F. Benavente, J. Barbosa, and V. Sanz-Nebot, *Analysis of recombinant human erythropoietin glycopeptides by capillary electrophoresis electrospray–time of flight-mass spectrometry,* Anal Chim Acta **709**, 81 (2012).

[115] D. Kolarich, P. H. Jensen, F. Altmann, and N. H. Packer, *Determination of site-specific glycan heterogeneity on glycoproteins,* Nat Protoc **7**, 1285 (2012).

[116] E. Vandermarliere, M. Mueller, and L. Martens, *Getting intimate with trypsin, the leading protease in proteomics,* Mass Spectrom Rev **32**, 453 (2013).

[117] J. Jiang, F. Tian, Y. Cai, X. Qian, C. E. Costello, and W. Ying, *Site-specific qualitative and quantitative analysis of the n-and o-glycoforms in recombinant human erythropoietin,* Anal Bioanal Chem **406**, 6265 (2014).

[118] J. M. Chalker, G. J. Bernardes, Y. A. Lin, and B. G. Davis, *Chemical modification of proteins at cysteine: opportunities in chemistry and biology,* Chem Asian J **4**, 630 (2009).

[119] H. Linley, *A new synthetic substrate for trypsin and its application to the determination of the amino-acid sequence of proteins,* Nature **178**, 647 (1956).

[120] A. Serra, X. Hemu, G. K. Nguyen, N. T. Nguyen, S. K. Sze, and J. P. Tam, *A high-throughput peptidomic strategy to decipher the molecular diversity of cyclic cysteine-rich peptides,* Sci Rep **6**, 23005 (2016).

[121] N. DeGraan-Weber and J. P. Reilly, *Use of cysteine aminoethylation to identify the hypervariable peptides of an antibody,* Anal Chem **90**, 1608 (2018).

[122] A. Büttner, M. Maier, L. Bonnington, P. Bulau, and D. Reusch, *Multi-attribute monitoring of complex erythropoetin beta glycosylation by gluc liquid chromatography mass spectrometry peptide mapping,* Anal Chem (2020).

[123] B. MacLean, D. M. Tomazela, N. Shulman, M. Chambers, G. L. Finney, B. Frewen, R. Kern, D. L. Tabb, D. C. Liebler, and M. J. MacCoss, *Skyline: an open source document editor for creating and analyzing targeted proteomics experiments,* Bioinformatics **26**, 966 (2010).

[124] E. S. Boja and H. M. Fales, *Overalkylation of a protein digest with iodoacetamide,* Anal Chem **73**, 3576 (2001).

[125] P. Liigand, K. Kaupmees, and A. Kruve, *Influence of the amino acid composition on the ionization efficiencies of small peptides,* J Mass Spectrom **54**, 481 (2019).

[126] K.-E. Song, J. Byeon, D.-B. Moon, H.-H. Kim, Y.-J. Choi, and J.-K. Suh, *Structural identification of modified amino acids on the interface between epo and its receptor from epo brp, human recombinant erythropoietin by lc/ms analysis,* Mol Cells **37**, 819 (2014).

[127] C. Winearls, M. Pippard, M. Downing, D. Oliver, C. Reid, and P. Mary Cotes, *Effect of human erythropoietin derived from recombinant dna on the anaemia of patients maintained by chronic haemodialysis,* The Lancet **328**, 1175 (1986).

[128] W. Jelkmann, *Erythropoietin: structure, control of production, and function,* Physiol. Rev. **72**, 449 (1992).

[129] T. B. Drüeke, F. Locatelli, N. Clyne, K.-U. Eckardt, I. C. Macdougall, D. Tsakiris, H.-U. Burger, A. Scherhag, and C. Investigators, *Normalization of hemoglobin level in patients with chronic kidney disease and anemia,* N Engl J Med **355**, 2071 (2006).

[130] M. Nimtz, W. Martin, V. Wray, K.-D. Klöppel, J. Augustin, and H. S. Conradt, *Structures of sialylated oligosaccharides of human erythropoietin expressed in recombinant bhk-21 cells,* Eur J Biochem **213**, 39 (1993).

[131] W. R. Alley, B. F. Mann, and M. V. Novotny, *High-sensitivity analytical approaches for the structural characterization of glycoproteins,* Chem Rev **113**, 2668 (2013).

[132] M. Girard, A. Puerta, J. C. Diez-Masa, and M. de Frutos, *High resolution separation methods for the determination of intact human erythropoiesis stimulating agents. a review,* Anal Chim Acta **713**, 7 (2012).

[133] E. Balaguer and C. Neusüss, *Glycoprotein characterization combining intact protein and glycan analysis by capillary electrophoresis-electrospray ionization-mass spectrometry,* Anal Chem **78**, 5384 (2006).

[134] R. G. Jayo, M. Thaysen-Andersen, P. W. Lindenburg, R. Haselberg, T. Hankemeier, R. Ramautar, and D. D. Y. Chen, *Simple capillary electrophoresis–mass spectrometry method for complex glycan analysis using a flow-through microvial interface,* Anal Chem **86**, 6479 (2014).

[135] A. Shubhakar, R. P. Kozak, K. R. Reiding, L. Royle, D. I. R. Spencer, D. L. Fernandes, and M. Wuhrer, *Automated high-throughput permethylation for glycosylation analysis of biologics using maldi-tof-ms,* Anal Chem **88**, 8562 (2016).

[136] P. Capdeville, L. Martin, S. Cholet, A. Damont, M. Audran, M. Ericsson, F. Fenaille, and A. Marchand, *Evaluation of erythropoietin biosimilars epotin™, hemax® and jimaixin™ by electrophoretic methods used for doping control analysis and specific n-glycan analysis revealed structural differences from original epoetin alfa drug eprex®,* J Pharm Biomed Anal **194**, 113750 (2021).

[137] H. Sasaki, N. Ochi, A. Dell, and M. Fukuda, *Site-specific glycosylation of human recombinant erythropoietin: analysis of glycopeptides or peptides at each glycosylation site by fast atom bombardment-mass spectrometry,* Biochemistry **27**, 8618 (1988).

[138] R. S. Rush, P. L. Derby, D. M. Smith, C. Merry, G. Rogers, M. F. Rohde, and V. Katta, *Microheterogeneity of erythropoietin carbohydrate structure,* Anal Chem **67**, 1442 (1995).

[139] N. Kawasaki, M. Ohta, S. Hyuga, M. Hyuga, and T. Hayakawa, *Application of liquid chromatography/mass spectrometry and liquid chromatography with tandem mass spectrometry to the analysis of the site-specific carbohydrate heterogeneity in erythropoietin,* Anal Biochem **285**, 82 (2000).

[140] S. Lippold, A. Büttner, M. S. F. Choo, M. Hook, C. J. de Jong, T. Nguyen-Khuong, M. Haberger, D. Reusch, M. Wuhrer, and N. de Haan, *Cysteine aminoethylation enables the site-specific glycosylation analysis of recombinant human erythropoietin using trypsin,* Anal Chem **92**, 9476 (2020).

[141] E. Watson, A. Bhide, and H. van Halbeek, *Structure determination of the intact major sialylated oligosaccharide chains of recombinant human erythropoietin expressed in chinese hamster ovary cells,* Glycobiology **4**, 227 (1994).

[142] C. Neusüß, U. Demelbauer, and M. Pelzing, *Glycoform characterization of intact erythropoietin by capillary electrophoresis-electrospray-time of flight-mass spectrometry,* Electrophoresis **26**, 1442 (2005).

[143] E. Giménez, F. Benavente, J. Barbosa, and V. Sanz-Nebot, *Analysis of intact erythropoietin and novel erythropoiesis-stimulating protein by capillary electrophoresis-electrospray-ion trap mass spectrometry,* Electrophoresis **29**, 2161 (2008).

[144] A. Pantazaki, M. Taverna, and C. Vidal-Madjar, *Recent advances in the capillary electrophoresis of recombinant glycoproteins,* Anal Chim Acta **383**, 137 (1999).

[145] E. Domínguez-Vega, S. Tengattini, C. Peintner, J. van Angeren, C. Temporini, R. Haselberg, G. Massolini, and G. W. Somsen, *High-resolution glycoform profiling of intact therapeutic proteins by hydrophilic interaction chromatography-mass spectrometry,* Talanta **184**, 375 (2018).

[146] A. Harazono, N. Hashii, R. Kuribayashi, S. Nakazawa, and N. Kawasaki, *Mass spectrometric glycoform profiling of the innovator and biosimilar erythropoietin and darbepoetin by lc/esi-ms,* J Pharm Biomed Anal **83**, 65 (2013).

[147] R. Haselberg, G. J. de Jong, and G. W. Somsen, *Low-flow sheathless capillary electrophoresis–mass spectrometry for sensitive glycoform profiling of intact pharmaceutical proteins,* Anal Chem **85**, 2289 (2013).

[148] A. Taichrib, M. Pelzing, C. Pellegrino, M. Rossi, and C. Neusüß, *High resolution tof ms coupled to ce for the analysis of isotopically resolved intact proteins,* J Proteomics **74**, 958 (2011).

[149] S. M. R. Stanley and A. Poljak, *Matrix-assisted laser-desorption time-of flight ionisation and high-performance liquid chromatography–electrospray ionisation mass spectral analyses of two glycosylated recombinant epoetins,* J Chromatogr B **785**, 205 (2003).

[150] E. Giménez, F. Benavente, J. Barbosa, and V. Sanz-Nebot, *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 (2007).

[151] E. Llop, R. Gutiérrez-Gallego, J. Segura, J. Mallorquí, and J. A. Pascual, *Structural analysis of the glycosylation of gene-activated erythropoietin (epoetin delta, dynepo),* Anal Biochem **383**, 243 (2008).

[152] E. Giménez, F. Benavente, J. Barbosa, and V. Sanz-Nebot, *Ionic liquid matrices for maldi-tof-ms analysis of intact glycoproteins,* Anal Bioanal Chem **398**, 357 (2010).

[153] X. Lou, J. L. J. van Dongen, and E. W. Meijer, *Generation of csi cluster ions for mass calibration in matrix-assisted laser desorption/ionization mass spectrometry,* J Am Soc Mass Spectrom **21**, 1223 (2010).

[154] M. Strohalm, M. Hassman, B. Košata, and M. Kodíček, *mmass data miner: an open source alternative for mass spectrometric data analysis,* Rapid Commun Mass Spectrom **22**, 905 (2008).

[155] J. Soltwisch, J. Souady, S. Berkenkamp, and K. Dreisewerd, *Effect of gas pressure and gas type on the fragmentation of peptide and oligosaccharide ions generated in an elevated pressure uv/ir-maldi ion source coupled to an orthogonal time-of-flight mass spectrometer,* Anal Chem **81**, 2921 (2009).

[156] D. Asakawa, D. Calligaris, T. A. Zimmerman, and E. De Pauw, *In-source decay during matrix-assisted laser desorption/ionization combined with the collisional process in an fticr mass spectrometer,* Anal Chem **85**, 7809 (2013).

[157] M. H. J. Selman, L. A. McDonnell, M. Palmblad, L. R. Ruhaak, A. M. Deelder, and M. Wuhrer, *Immunoglobulin g glycopeptide profiling by matrix-assisted laser desorption ionization fourier transform ion cyclotron resonance mass spectrometry,* Anal Chem **82**, 1073 (2010).

[158] K. R. Reiding, L. R. Ruhaak, H.-W. Uh, S. el Bouhaddani, E. B. van den Akker, R. Plomp, L. A. McDonnell, J. J. Houwing-Duistermaat, P. E. Slagboom, M. Beekman, and M. Wuhrer, *Human plasma n-glycosylation as analyzed by matrix-assisted laser desorption/ionization-fourier transform ion cyclotron resonance-ms associates with markers of inflammation and metabolic health,* Mol Cell Proteom **16**, 228 (2017).

[159] N. de Haan, S. Yang, J. Cipollo, and M. Wuhrer, *Glycomics studies using sialic acid derivatization and mass spectrometry,* Nat Rev Chem **4**, 229 (2020).

[160] N. de Haan, K. R. Reiding, M. Haberger, D. Reusch, D. Falck, and M. Wuhrer, *Linkage-specific sialic acid derivatization for maldi-tof-ms profiling of igg glycopeptides,* Anal Chem **87**, 8284 (2015).

[161] H. Li, W. Gao, X. Feng, B.-F. Liu, and X. Liu, *Maldi-ms analysis of sialylated n-glycan linkage isomers using solid-phase two step derivatization method,* Anal Chim Acta **924**, 77 (2016).

[162] J. J. Pitt and J. J. Gorman, *Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of sialylated glycopeptides and proteins using 2,6-dihydroxyacetophenone as a matrix,* Rapid Commun Mass Spectrom **10**, 1786 (1996).

[163] J. A. Castoro and C. L. Wilkins, *Ultrahigh resolution matrix-assisted laser desorption/ionization of small proteins by fourier-transform mass spectrometry,* Anal Chem **65**, 2621 (1993).

[164] M. L. Easterling, C. C. Pitsenberger, S. S. Kulkarni, P. K. Taylor, and I. J. Amster, *A 4.7 tesla internal maldi-fticr instrument for high mass studies: performance and methods,* Int J Mass Spectrom Ion Processes **157-158**, 97 (1996).

[165] A. G. Marshall, C. L. Hendrickson, and G. S. Jackson, *Fourier transform ion cyclotron resonance mass spectrometry: A primer,* Mass Spectrom Rev **17**, 1 (1998).

[166] H.-Q. Nguyen, D. Lee, Y. Kim, M. Paek, M. Kim, K.-S. Jang, J. Oh, Y.-S. Lee, J. E. Yeon, D. M. Lubman, and J. Kim, *Platelet factor 4 as a novel exosome marker in maldi-ms analysis of exosomes from human serum,* Anal Chem **91**, 13297 (2019).

[167] F. Fleurbaaij, M. E. M. Kraakman, E. C. J. Claas, C. W. Knetsch, H. C. van Leeuwen, Y. E. M. van der Burgt, K. E. Veldkamp, M. C. Vos, W. Goessens, B. J. Mertens, E. J. Kuijper, P. J. Hensbergen, and S. Nicolardi, *Typing pseudomonas aeruginosa isolates with ultrahigh resolution maldi-fticr mass spectrometry,* Anal Chem **88**, 5996 (2016).

[168] J. M. Spraggins, D. G. Rizzo, J. L. Moore, M. J. Noto, E. P. Skaar, and R. M. Caprioli, *Next-generation technologies for spatial proteomics: Integrating ultra-high speed maldi-tof and high mass resolution maldi fticr imaging mass spectrometry for protein analysis,* Proteomics **16**, 1678 (2016).

[169] B. M. Prentice, D. J. Ryan, R. Van de Plas, R. M. Caprioli, and J. M. Spraggins, *Enhanced ion transmission efficiency up to m/z 24 000 for maldi protein imaging mass spectrometry,* Anal Chem **90**, 5090 (2018).

[170] S. Nicolardi, L. Switzar, A. M. Deelder, M. Palmblad, and Y. E. M. van der Burgt, *Top-down maldi-in-source decay-fticr mass spectrometry of isotopically resolved proteins,* Anal Chem **87**, 3429 (2015).

[171] S. Nicolardi, Y. E. M. van der Burgt, M. Wuhrer, and A. M. Deelder, *Mapping o-glycosylation of apolipoprotein c-iii in maldi-ft-icr protein profiles,* Proteomics **13**, 992 (2013).

[172] Y. E. M. van der Burgt, D. P. A. Kilgour, Y. O. Tsybin, K. Srzentić, L. Fornelli, A. Beck, M. Wuhrer, and S. Nicolardi, *Structural analysis of monoclonal antibodies by ultrahigh resolution maldi in-source decay ft-icr mass spectrometry,* Anal Chem **91**, 2079 (2019).

[173] C. Gstöttner, D. Reusch, M. Haberger, I. Dragan, P. Van Veelen, D. P. A. Kilgour, Y. O. Tsybin, Y. E. M. van der Burgt, M. Wuhrer, and S. Nicolardi, *Monitoring glycation levels of a bispecific monoclonal antibody at subunit level by ultrahigh-resolution maldi ft-icr mass spectrometry,* mAbs **12**, 1682403 (2020).

[174] S. Nicolardi, A. A. Joseph, Q. Zhu, Z. Shen, A. Pardo-Vargas, F. Chiodo, A. Molinaro, A. Silipo, Y. E. van der Burgt, B. Yu, P. H. Seeberger, and M. Wuhrer, *Analysis of synthetic monodisperse polysaccharides by wide mass range ultrahigh-resolution maldi mass spectrometry,* Anal Chem , 10.1021/acs.analchem.1c00239, In Press (2021).

[175] Q. Zhu, Z. Shen, F. Chiodo, S. Nicolardi, A. Molinaro, A. Silipo, and B. Yu, *Chemical synthesis of glycans up to a 128-mer relevant to the o-antigen of bacteroides vulgatus,* Nat Commun **11**, 4142 (2020).

[176] G. Stübiger, M. Marchetti, M. Nagano, C. Reichel, G. Gmeiner, and G. Allmaier, *Characterisation of intact recombinant human erythropoietins applied in doping by means of planar gel electrophoretic techniques and matrix-assisted laser desorption/ionisation linear time-of-flight mass spectrometry,* Rapid Commun Mass Spectrom **19**, 728 (2005).

[177] X.-R. Jiang, A. Song, S. Bergelson, T. Arroll, B. Parekh, K. May, S. Chung, R. Strouse, A. Mire-Sluis, and M. Schenerman, *Advances in the assessment and control of the effector functions of therapeutic antibodies,* Nat Rev Drug Discov **10**, 101–111 (2011).

[178] J. B. Evans and B. A. Syed, *Next-generation antibodies,* Nat Rev Drug Discov **13**, 413–414 (2014).

[179] F. Cymer, H. Beck, A. Rohde, and D. Reusch, *Therapeutic monoclonal antibody n-glycosylation–structure, function and therapeutic potential,* Biologicals **52**, 1 (2018).

[180] P. Sondermann, R. Huber, V. Oosthuizen, and U. Jacob, *The 3.2-Å crystal structure of the human igg1 fc fragment–fcγriii complex,* Nature **406**, 267 (2000).

[181] T. Dashivets, M. Thomann, P. Rueger, A. Knaupp, J. Buchner, and T. Schlothauer, *Multi-angle effector function analysis of human monoclonal igg glycovariants,* PLoS One **10**, e0143520 (2015).

[182] H. Li, H. H. Nguyen, R. R. O. Loo, I. D. Campuzano, and J. A. Loo, *An integrated native mass spectrometry and top-down proteomics method that connects sequence to structure and function of macromolecular complexes,* Nat Chem **10**, 139–148 (2018).

[183] S. Rosati, Y. Yang, A. Barendregt, and A. J. R. Heck, *Detailed mass analysis of structural heterogeneity in monoclonal antibodies using native mass spectrometry,* Nat Protoc **9**, 967–976 (2014).

[184] A. O. Bailey, G. Han, W. Phung, P. Gazis, J. Sutton, J. L. Josephs, and W. Sandoval, *Charge variant native mass spectrometry benefits mass precision and dynamic range of monoclonal antibody intact mass analysis,* mAbs **10**, 1214 (2018).

[185] I. D. Campuzano, H. Li, D. Bagal, J. L. Lippens, J. Svitel, R. J. Kurzeja, H. Xu, P. D. Schnier, and J. A. Loo, *Native ms analysis of bacteriorhodopsin and an empty nanodisc by orthogonal acceleration time-of-flight, orbitrap and ion cyclotron resonance,* Anal Chem **88**, 12427 (2016).

[186] M. Kiyoshi, J. M. Caaveiro, M. Tada, H. Tamura, T. Tanaka, Y. Terao, K. Morante, A. Harazono, N. Hashii, and H. Shibata, *Assessing the heterogeneity of the fc-glycan of a therapeutic antibody using an engineered fcγreceptor iiia-immobilized column,* Sci Rep **8**, 3955 (2018).

[187] K. R. Patel, J. T. Roberts, G. P. Subedi, and A. W. Barb, *Restricted processing of cd16a/fc γ receptor iiia n-glycans from primary human nk cells impacts structure and function,* J Biol Chem **293**, 3477–3489 (2018).

[188] N. Washburn, R. Meccariello, J. Duffner, K. Getchell, K. Holte, T. Prod'homme, K. Srinivasan, R. Prenovitz, J. Lansing, and I. Capila, *Characterization of endogenous human fcγriii by mass spectrometry reveals site, allele and sequence specific glycosylation,* Mol Cell Proteomics **18**, 534 (2019).

[189] A. Zeck, G. Pohlentz, T. Schlothauer, J. Peter-Katalinic, and J. T. Regula, *Cell type-specific and site directed n-glycosylation pattern of fcγriiia,* J. Proteome Res. **10**, 3031 (2011).

[190] C. Ferrara, P. Brünker, T. Suter, S. Moser, U. Püntener, and P. Umaña, *The carbohydrate at fcγriiia asn-162 an element required for high affinity binding to non-fucosylated igg glycoforms,* Biotechnol. Bioeng. **93**, 851–861 (2006).

[191] G. Dekkers, R. Plomp, C. A. Koeleman, R. Visser, H. H. Von Horsten, V. Sandig, T. Rispens, M. Wuhrer, and G. Vidarsson, *Multi-level glyco-engineering techniques to generate igg with defined fc-glycans,* Sci Rep **6**, 36964 (2016).

[192] M. Yu, D. Brown, C. Reed, S. Chung, J. Lutman, E. Stefanich, A. Wong, J.-P. Stephan, and R. Bayer, *Production, characterization and pharmacokinetic properties of antibodies with n-linked mannose-5 glycans,* mAbs **4**, 475–487 (2012).

[193] Y. D. Liu and G. C. Flynn, *Effect of high mannose glycan pairing on igg antibody clearance,* Biologicals **44**, 163–169 (2016).

[194] Y. Yang, G. Wang, T. Song, C. B. Lebrilla, and A. J. Heck, *Resolving the micro-heterogeneity and structural integrity of monoclonal antibodies by hybrid mass spectrometric approaches,* mAbs **9**, 638 (2017).

[195] S. Ha, Y. Ou, J. Vlasak, Y. Li, S. Wang, K. Vo, Y. Du, A. Mach, Y. Fang, and N. Zhang, *Isolation and characterization of igg1 with asymmetrical fc glycosylation,* Glycobiology **21**, 1087 (2011).

[196] A. Pincetic, S. Bournazos, D. J. DiLillo, J. Maamary, T. T. Wang, R. Dahan, B. M. Fiebiger, and J. V. Ravetch, *Type i and type ii fc receptors regulate innate and adaptive immunity,* Nat Immunol **15**, 707 (2014).

[197] K. O. Saunders, *Conceptual approaches to modulating antibody effector functions and circulation half-life,* Front Immunol **10** (2019).

[198] S. Lippold, S. Nicolardi, E. Domínguez-Vega, A.-K. Heidenreich, G. Vidarsson, D. Reusch, M. Haberger, M. Wuhrer, and D. Falck, *Glycoform-resolved fcγriiia affinity chromatography–mass spectrometry,* mAbs , 1191 (2019).

[199] A. C. Chan and P. J. Carter, *Therapeutic antibodies for autoimmunity and inflammation,* Nat Rev Immunol **10**, 301 (2010).

[200] L. M. Weiner, R. Surana, and S. Wang, *Monoclonal antibodies: versatile platforms for cancer immunotherapy,* Nat Rev Immunol **10**, 317 (2010).

[201] C. Klein, W. Schaefer, and J. T. Regula, *The use of crossmab technology for the generation of bi- and multispecific antibodies,* mAbs **8**, 1010 (2016).

[202] K. R. Anumula, *Quantitative glycan profiling of normal human plasma derived immunoglobulin and its fragments fab and fc,* J Immunol Methods **382**, 167 (2012).

[203] J. Kurai, H. Chikumi, K. Hashimoto, K. Yamaguchi, A. Yamasaki, T. Sako, H. Touge, H. Makino, M. Takata, M. Miyata, M. Nakamoto, N. Burioka, and E. Shimizu, *Antibody-dependent cellular cytotoxicity mediated by cetuximab against lung cancer cell lines,* Clin Cancer Res **13**, 1552 (2007).

[204] A. Kol, A. Terwisscha van Scheltinga, M. Pool, C. Gerdes, E. de Vries, and S. de Jong, *Adcc responses and blocking of egfr-mediated signaling and cell growth by combining the anti-egfr antibodies imgatuzumab and cetuximab in nsclc cells,* Oncotarget **8**, 45432 (2017).

[205] M. Bern, T. Caval, Y. J. Kil, W. Tang, C. Becker, E. Carlson, D. Kletter, K. I. Sen, N. Galy, D. Hagemans, V. Franc, and A. J. R. Heck, *Parsimonious charge deconvolution for native mass spectrometry,* J Proteome Res **17**, 1216 (2018).

[206] F. Lermyte, Y. O. Tsybin, P. B. O'Connor, and J. A. Loo, *Top or middle? up or down? toward a standard lexicon for protein top-down and allied mass spectrometry approaches,* J Am Soc Mass Spectrom **30**, 1149 (2019).

[207] Z. Zhang, R. Perrault, Y. Zhao, and J. Ding, *Speb proteolysis with imaged capillary isoelectric focusing for the characterization of domain-specific charge heterogeneities of reference and biosimilar rituximab,* J Chromatogr B **1020**, 148 (2016).

[208] M. C. Janin-Bussat, L. Tonini, C. Huillet, O. Colas, C. Klinguer-Hamour, N. Corvaia, and A. Beck, *Cetuximab fab and fc n-glycan fast characterization using ides digestion and liquid chromatography coupled to electrospray ionization mass spectrometry,* Methods Mol Biol **988**, 93 (2013).

[209] T. Schlothauer, S. Herter, C. F. Koller, S. Grau-Richards, V. Steinhart, C. Spick, M. Kubbies, C. Klein, P. Umana, and E. Mossner, *Novel human igg1 and igg4 fc-engineered antibodies with completely abolished immune effector functions,* Protein Eng. Des. Sel. **29**, 457 (2016).

[210] B. Yan, D. Boyd, T. Kaschak, J. Tsukuda, A. Shen, Y. Lin, S. Chung, P. Gupta, A. Kamath, A. Wong, J. M. Vernes, G. Y. Meng, K. Totpal, G. Schaefer, G. Jiang, B. Nogal, C. Emery, M. Vanderlaan, P. Carter, R. Harris, and A. Amanullah, *Engineering upper hinge improves stability and effector function of a human igg1,* J Biol Chem **287**, 5891 (2012).

[211] H. Liu and K. May, *Disulfide bond structures of igg molecules: structural variations, chemical modifications and possible impacts to stability and biological function,* mAbs **4**, 17 (2012).

[212] H. S. Kim, I. Kim, L. Zheng, J.-M. Vernes, Y. G. Meng, and C. Spiess, *Evading pre-existing anti-hinge antibody binding by hinge engineering,* mAbs **8**, 1536 (2016).

[213] E. T. J. van den Bremer, A. F. Labrijn, R. van den Boogaard, P. Priem, K. Scheffler, J. P. M. Melis, J. Schuurman, P. Parren, and R. N. de Jong, *Cysteine-silac mass spectrometry enabling the identification and quantitation of scrambled interchain disulfide bonds: Preservation of native heavy-light chain pairing in bispecific iggs generated by controlled fab-arm exchange,* Anal Chem **89**, 10873 (2017).

[214] T. S. Raju and B. Scallon, *Fc glycans terminated with n-acetylglucosamine residues increase antibody resistance to papain,* Biotechnol Prog **23**, 964 (2007).

[215] M. T. Tyn and T. W. Gusek, *Prediction of diffusion coefficients of proteins,* Biotechnol Bioeng **35**, 327 (1990).

[216] F. Gritti and G. Guiochon, *Mass transfer kinetics, band broadening and column efficiency,* J Chromatogr A **1221**, 2 (2012).

[217] A. Astefanei, I. Dapic, and M. Camenzuli, *Different stationary phase selectivities and morphologies for intact protein separations,* Chromatographia **80**, 665 (2017).

[218] M. Aoyama, N. Hashii, W. Tsukimura, K. Osumi, A. Harazono, M. Tada, M. Kiyoshi, A. Matsuda, and A. Ishii-Watabe, *Effects of terminal galactose residues in mannose α 1-6 arm of fc-glycan on the effector functions of therapeutic monoclonal antibodies,* mAbs **11**, 826 (2019).

[219] M. Biacchi, R. Gahoual, N. Said, A. Beck, E. Leize-Wagner, and Y. N. Francois, *Glycoform separation and characterization of cetuximab variants by middle-up off-line capillary zone electrophoresis-uv/electrospray ionization-ms,* Anal Chem **87**, 6240 (2015).

[220] I. D. G. Campuzano, J. H. Robinson, J. O. Hui, S. D. H. Shi, C. Netirojjanakul, M. Nshanian, P. F. Egea, J. L. Lippens, D. Bagal, J. A. Loo, and M. Bern, *Native and denaturing ms protein deconvolution for biopharma: Monoclonal antibodies and antibody–drug conjugates to polydisperse membrane proteins and beyond,* Anal Chem , in press (2019).

[221] K. Brorson and A. Y. Jia, *Therapeutic monoclonal antibodies and consistent ends: terminal heterogeneity, detection, and impact on quality,* Curr Opin Biotechnol **30**, 140 (2014).

[222] E. T. J. van den Bremer, F. J. Beurskens, M. Voorhorst, P. J. Engelberts, R. N. de Jong, B. G. van der Boom, E. M. Cook, M. A. Lindorfer, R. P. Taylor, P. H. C. van Berkel, and P. W. H. I. Parren, *Human igg is produced in a pro-form that requires clipping of c-terminal lysines for maximal complement activation,* mAbs **7**, 672 (2015).

[223] L. Huang, S. Biolsi, K. R. Bales, and U. Kuchibhotla, *Impact of variable domain glycosylation on antibody clearance: an lc/ms characterization,* Anal Biochem **349**, 197 (2006).

[224] R. Jefferis, *Antibody therapeutics: isotype and glycoform selection,* Expert Opin Biol Ther **7**, 1401 (2007).

[225] A. Wiegandt and B. Meyer, *Unambiguous characterization of n-glycans of monoclonal antibody cetuximab by integration of lc-ms/ms and 1h nmr spectroscopy,* Anal Chem **86**, 4807 (2014).

[226] S. Radaev, S. Motyka, W.-H. Fridman, C. Sautes-Fridman, and P. D. Sun, *The structure of a human type iii fcγ receptor in complex with fc*,* J Biol Chem **276**, 16469 (2001).

[227] A. W. Barb, *Fc γ receptor compositional heterogeneity: Considerations for immunotherapy development,* J Biol Chem **296**, 100057 (2021).

[228] C. Kerntke, F. Nimmerjahn, and M. Biburger, *There is (scientific) strength in numbers: A comprehensive quantitation of fc gamma receptor numbers on human and murine peripheral blood leukocytes,* Front Immunol **11** (2020).

[229] G. P. Subedi and A. W. Barb, *The immunoglobulin g1 n-glycan composition affects binding to each low affinity fc γ receptor,* mAbs **8**, 1512 (2016).

[230] J. T. Roberts and A. W. Barb, *A single amino acid distorts the fc γ receptor iiib/cd16b structure upon binding immunoglobulin g1 and reduces affinity relative to cd16a,* J Biol Chem **293**, 19899 (2018).

[231] J. M. Caaveiro, M. Kiyoshi, and K. Tsumoto, *Structural analysis of fc/fcγr complexes: a blueprint for antibody design,* Immunol Rev **268**, 201 (2015).

[232] P. Bruhns, B. Iannascoli, P. England, D. A. Mancardi, N. Fernandez, S. Jorieux, and M. Daëron, *Specificity and affinity of human fcγ receptors and their polymorphic variants for human igg subclasses,* Blood **113**, 3716 (2009).

[233] M. Peipp, J. J. Lammerts van Bueren, T. Schneider-Merck, W. W. K. Bleeker, M. Dechant, T. Beyer, R. Repp, P. H. C. van Berkel, T. Vink, J. G. J. van de Winkel, P. W. H. I. Parren, and T. Valerius, *Antibody fucosylation differentially impacts cytotoxicity mediated by nk and pmn effector cells,* Blood **112**, 2390 (2008).

[234] G. P. Subedi and A. W. Barb, *Cd16a with oligomannose-type n-glycans is the only "low-affinity" fc γ receptor that binds the igg crystallizable fragment with high affinity in vitro,* J Biol Chem **293**, 16842 (2018).

[235] S. Lippold, S. Nicolardi, M. Wuhrer, and D. Falck, *Proteoform-resolved fcγriiia binding assay for fab glycosylated monoclonal antibodies achieved by affinity chromatography mass spectrometry of fc moieties,* Front Chem **7** (2019).

[236] I. Wojcik, T. Sénard, E. L. de Graaf, G. M. C. Janssen, A. H. de Ru, Y. Mohammed, P. A. van Veelen, G. Vidarsson, M. Wuhrer, and D. Falck, *Site-specific glycosylation mapping of fc gamma receptor iiib from neutrophils of individual healthy donors,* Anal Chem **92**, 13172 (2020).

[237] P. Umaña, J. Jean–Mairet, R. Moudry, H. Amstutz, and J. E. Bailey, *Engineered glycoforms of an antineuroblastoma igg1 with optimized antibody-dependent cellular cytotoxic activity,* Nat Biotech **17**, 176 (1999).

[238] J. M. Hayes, A. Frostell, R. Karlsson, S. Müller, S. M. Martín, M. Pauers, F. Reuss, E. F. Cosgrave, C. Anneren, G. P. Davey, and P. M. Rudd, *Identification of fc gamma receptor glycoforms that produce differential binding kinetics for rituximab*,* Mol Cell Proteom **16**, 1770 (2017).

[239] F. Cambay, C. Forest-Nault, L. Dumoulin, A. Seguin, O. Henry, Y. Durocher, and G. De Crescenzo, *Glycosylation of fcγ receptors influences their interaction with various igg1 glycoforms,* Mol Immunol **121**, 144 (2020).

[240] T. L. J. van Osch, J. Nouta, N. I. L. Derksen, G. van Mierlo, C. E. van der Schoot, M. Wuhrer, T. Rispens, and G. Vidarsson, *Fc galactosylation promotes hexamerization of human igg1, leading to enhanced classical complement activation,* J Immunol (2021).

[241] B. Wei, X. Gao, L. Cadang, S. Izadi, P. Liu, H.-M. Zhang, E. Hecht, J. Shim, G. Magill, J. R. Pabon, L. Dai, W. Phung, E. Lin, C. Wang, K. Whang,

S. Sanchez, J. Oropeza Jr, J. Camperi, J. Zhang, W. Sandoval, Y. T. Zhang, and G. Jiang, *Fc galactosylation follows consecutive reaction kinetics and enhances immunoglobulin g hexamerization for complement activation,* mAbs **13**, 1893427 (2021).

[242] R. Yogo, Y. Yamaguchi, H. Watanabe, H. Yagi, T. Satoh, M. Nakanishi, M. Onitsuka, T. Omasa, M. Shimada, T. Maruno, T. Torisu, S. Watanabe, D. Higo, T. Uchihashi, S. Yanaka, S. Uchiyama, and K. Kato, *The fab portion of immunoglobulin g contributes to its binding to fcγ receptor iii,* Sci Rep **9**, 11957 (2019).

[243] Y. Sun, S. Izadi, M. Callahan, G. Deperalta, and A. T. Wecksler, *Antibody–receptor interactions mediate antibody-dependent cellular cytotoxicity,* J Biol Chem **297**, 100826 (2021).

[244] A. R. Evans, M. T. Capaldi, G. Goparaju, D. Colter, F. F. Shi, S. Aubert, L.-C. Li, J. Mo, M. J. Lewis, P. Hu, P. Alfonso, and P. Mehndiratta, *Using bispecific antibodies in forced degradation studies to analyze the structure–function relationships of symmetrically and asymmetrically modified antibodies,* mAbs **11**, 1101 (2019).

[245] C. Bahou, E. A. Love, S. Leonard, R. J. Spears, A. Maruani, K. Armour, J. R. Baker, and V. Chudasama, *Disulfide modified igg1: An investigation of biophysical profile and clinically relevant fc interactions,* Bioconjug Chem **30**, 1048 (2019).

[246] S. Derer, P. Glorius, M. Schlaeth, S. Lohse, K. Klausz, U. Muchhal, J. R. Desjarlais, A. Humpe, T. Valerius, and M. Peipp, *Increasing fcγriia affinity of an fcγriii-optimized anti-egfr antibody restores neutrophil-mediated cytotoxicity,* mAbs **6**, 409 (2014).

[247] L. W. Treffers, M. van Houdt, C. W. Bruggeman, M. H. Heineke, X. W. Zhao, J. van der Heijden, S. Q. Nagelkerke, P. J. J. H. Verkuijlen, J. Geissler, S. Lissenberg-Thunnissen, T. Valerius, M. Peipp, K. Franke, R. van Bruggen, T. W. Kuijpers, M. van Egmond, G. Vidarsson, H. L. Matlung, and T. K. van den Berg, *Fcγriiib restricts antibody-dependent destruction of cancer cells by human neutrophils,* Front Immunol **9**, 3124 (2019).

[248] E. García-García, G. Nieto-Castañeda, M. Ruiz-Saldaña, N. Mora, and C. Rosales, *Fcγriia and fcγriiib mediate nuclear factor activation through separate signaling pathways in human neutrophils,* J Immunol **182**, 4547 (2009).

[249] M. Guilliams, P. Bruhns, Y. Saeys, H. Hammad, and B. N. Lambrecht, *The function of fcγ receptors in dendritic cells and macrophages,* Nat Rev Immunol **14**, 94 (2014).

[250] F. Weber, D. Breustedt, S. Schlicht, C. A. Meyer, J. Niewoehner, M. Ebeling, P. O. Freskgard, P. Bruenker, T. Singer, M. Reth, and A. Iglesias, *First infusion reactions are mediated by fcγriiib and neutrophils,* Pharm Res **35**, 169 (2018).

[251] A. Freimoser–Grundschober, P. Rueger, F. Fingas, P. Sondermann, S. Herter, T. Schlothauer, P. Umana, and C. Neumann, *Fcγriiia chromatography to enrich a-fucosylated glycoforms and assess the potency of glycoengineered therapeutic antibodies,* J. Chromatogr. A **1610**, 460554 (2020).

[252] K. T. Pang, S. J. Tay, C. Wan, I. Walsh, M. S. F. Choo, Y. S. Yang, A. Choo, Y. S. Ho, and T. Nguyen-Khuong, *Semi-automated glycoproteomic data analysis of lc-ms data using glycopeptidegraphms in process development of monoclonal antibody biologics,* Front Chem **9** (2021).

[253] R. Kawahara, K. Alagesan, M. Bern, W. Cao, R. J. Chalkley, K. Cheng, M. S. Choo, N. Edwards, R. Goldman, M. Hoffmann, Y. Hu, Y. Huang, J. Y. Kim, D. Kletter, B. Liquet-Weiland, M. Liu, Y. Mechref, B. Meng, S. Neelamegham, T. Nguyen-Khuong, J. Nilsson, A. Pap, G. W. Park, B. L. Parker, C. L. Pegg, J. M. Penninger, T. K. Phung, M. Pioch, E. Rapp, E. Sakalli, M. Sanda, B. L. Schulz, N. E. Scott, G. Sofronov, J. Stadlmann, S. Y. Vakhrushev, C. M. Woo, H.-Y. Wu, P. Yang, W. Ying, H. Zhang, Y. Zhang, J. Zhao, J. Zaia, S. M. Haslam, G. Palmisano, J. S. Yoo, G. Larson, K.-H. Khoo, K. F. Medzihradszky, D. Kolarich, N. H. Packer, and M. Thaysen-Andersen, *Community evaluation of glycoproteomics informatics solutions reveals high-performance search strategies of glycopeptide data,* bioRxiv (2021).

[254] M. Hoffmann, M. Pioch, A. Pralow, R. Hennig, R. Kottler, U. Reichl, and E. Rapp, *The fine art of destruction: A guide to in-depth glycoproteomic analyses—exploiting the diagnostic potential of fragment ions,* Proteomics **18**, 1800282 (2018).

[255] W. Cao, M. Liu, S. Kong, M. Wu, Y. Zhang, and P. Yang, *Recent advances in software tools for more generic and precise intact glycopeptide analysis,* Mol Cell Proteomics , 100060 (2021).

[256] E. Ang, H. Neustaeter, V. Spicer, H. Perreault, and O. Krokhin, *Retention time prediction for glycopeptides in reversed-phase chromatography for glycoproteomic applications,* Anal Chem **91**, 13360 (2019).

[257] A. R. Evans, A. S. Hebert, J. Mulholland, M. J. Lewis, and P. Hu, *Id–mam: A validated identity and multi-attribute monitoring method for commercial release and stability testing of a bispecific antibody,* Anal Chem (2021).

[258] T. Robin, J. Mariethoz, and F. Lisacek, *Examining and fine-tuning the selection of glycan compositions with glyconnect compozitor,* Mol Cell Proteom **19**, 1602 (2020).

[259] A. Roushan, G. M. Wilson, D. Kletter, K. I. Sen, W. Tang, Y. J. Kil, E. Carlson, and M. Bern, *Peak filtering, peak annotation, and wildcard search for glycoproteomics,* Mol Cell Proteom **20**, 100011 (2021).

[260] A. C. Nichols, Y. J. Kil, A. Mahan, B. Zhai, R. Hepler, K. Nields, H. Nanda, E. Carlson, and M. Bern, *Orthogonal comparison of analytical methods by theoretical reconstruction from bottom-up assay data,* J Am Soc Mass Spectrom (2021).

[261] M. Thevis, R. R. Ogorzalek Loo, and J. A. Loo, *In-gel derivatization of proteins for cysteine-specific cleavages and their analysis by mass spectrometry,* J. Proteome Res. **2**, 163 (2003).

[262] T. Čaval, A. Buettner, M. Haberger, D. Reusch, and A. J. R. Heck, *Discrepancies between high-resolution native and glycopeptide-centric mass spectrometric approaches: A case study into the glycosylation of erythropoietin variants,* J Am Soc Mass Spectrom (2021).

[263] C. Gstöttner, T. Zhang, A. Resemann, S. Ruben, S. Pengelley, D. Suckau, T. Welsink, M. Wuhrer, and E. Domínguez-Vega, *Structural and functional characterization of sars-cov-2 rbd domains produced in mammalian cells,* Anal Chem **93**, 6839 (2021).

[264] T. Wohlschlager, K. Scheffler, I. C. Forstenlehner, W. Skala, S. Senn, E. Damoc, J. Holzmann, and C. G. Huber, *Native mass spectrometry combined with enzymatic dissection unravels glycoform heterogeneity of biopharmaceuticals,* Nat Commun **9**, 1713 (2018).

[265] M. T. Marty, *A universal score for deconvolution of intact protein and native electrospray mass spectra,* Anal Chem **92**, 4395 (2020).

[266] H. Choi, D. Lee, Y. Kim, H.-Q. Nguyen, S. Han, and J. Kim, *Effects of matrices and additives on multiple charge formation of proteins in maldi–ms analysis,* J Am Soc Mass Spectrom **30**, 1174 (2019).

[267] Y. Yan, T. Xing, S. Wang, and N. Li, *Versatile, sensitive, and robust native lc–ms platform for intact mass analysis of protein drugs,* J Am Soc Mass Spectrom **31**, 2171 (2020).

[268] L. Signor and E. Boeri Erba, *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).

[269] U. Leurs, U. H. Mistarz, and K. D. Rand, *Getting to the core of protein pharmaceuticals – comprehensive structure analysis by mass spectrometry,* Eur J Pharm Biopharm **93**, 95 (2015).

[270] F. Füssl, L. Strasser, S. Cari, and J. Bones, *Native lc–ms for capturing quality attributes of biopharmaceuticals on the intact protein level,* Curr. Opin. Biotechnol. **71**, 32 (2021).

[271] J. Camperi, V. Pichon, and N. Delaunay, *Separation methods hyphenated to mass spectrometry for the characterization of the protein glycosylation at the intact level,* J Pharm Biomed Anal **178**, 112921 (2020).

[272] F. Füssl, K. Cook, K. Scheffler, A. Farrell, S. Mittermayr, and J. Bones, *Charge variant analysis of monoclonal antibodies using direct coupled ph gradient cation exchange chromatography to high-resolution native mass spectrometry,* Anal Chem **90**, 4669 (2018).

[273] M. Haberger, K. Bomans, K. Diepold, M. Hook, J. Gassner, T. Schlothauer, A. Zwick, C. Spick, J. F. Kepert, B. Hienz, M. Wiedmann, H. Beck, P. Metzger, M. Mølhøj, C. Knoblich, U. Grauschopf, D. Reusch, and P. Bulau, *Assessment of chemical modifications of sites in the cdrs of recombinant antibodies,* mAbs **6**, 327 (2014).

[274] J. Camperi, A. Goyon, D. Guillarme, K. Zhang, and C. Stella, *Multi-dimensional lc-ms: the next generation characterization of antibody-based therapeutics by unified online bottom-up, middle-up and intact approaches,* Analyst **146**, 747 (2021).

[275] J. Camperi, L. Dai, D. Guillarme, and C. Stella, *Fast and automated characterization of monoclonal antibody minor variants from cell cultures by combined protein-a and multidimensional lc/ms methodologies,* Anal Chem **92**, 8506 (2020).

[276] J. A. Melby, D. S. Roberts, E. J. Larson, K. A. Brown, E. F. Bayne, S. Jin, and Y. Ge, *Novel strategies to address the challenges in top-down proteomics,* J Am Soc Mass Spectrom **32**, 1278 (2021).

[277] T. A. Millward, M. Heitzmann, K. Bill, U. Längle, P. Schumacher, and K. Forrer, *Effect of constant and variable domain glycosylation on pharmacokinetics of therapeutic antibodies in mice,* Biologicals **36**, 41 (2008).

[278] Z. Hao, T. Zhang, Y. Xuan, H. Wang, J. Qian, S. Lin, J. Chen, D. M. Horn, D. Argoti, A. Beck, S. Cianférani, P. Bennett, K. Miller, and A. Makarov, *Intact antibody characterization using orbitrap mass spectrometry,* in *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques*, ACS Symposium Series, Vol. 1202 (J. Am. Chem. Soc., 2015) Book section 10, pp. 289–315.

[279] I. Sokolowska, J. Mo, F. Rahimi Pirkolachahi, C. McVean, L. A. T. Meijer, L. Switzar, C. Balog, M. J. Lewis, and P. Hu, *Implementation of a high-resolution liquid chromatography–mass spectrometry method in quality control laboratories for release and stability testing of a commercial antibody product,* Anal Chem **92**, 2369 (2020).

[280] R. Karlsson, A. Michaelsson, and L. Mattsson, *Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system,* J. Immunol. Methods **145**, 229 (1991).

[281] S. N. Davidoff, N. T. Ditto, A. E. Brooks, J. Eckman, and B. D. Brooks, *Surface plasmon resonance for therapeutic antibody characterization,* in *Label-Free Biosensor Methods in Drug Discovery,* edited by Y. Fang (Springer New York, New York, NY, 2015) pp. 35–76.

[282] G. P. Jadhav, P. K. Prathipati, and H. Chauhan, *Surface plasmon resonance, orbitrap mass spectrometry and raman advancements: exciting new techniques in drug discovery,* Expert Opin Drug Deliv **15**, 739 (2020).

[283] D. Wassaf, G. Kuang, K. Kopacz, Q.-L. Wu, Q. Nguyen, M. Toews, J. Cosic, J. Jacques, S. Wiltshire, J. Lambert, C. C. Pazmany, S. Hogan, R. C. Ladner, A. E. Nixon, and D. J. Sexton, *High-throughput affinity ranking of antibodies using surface plasmon resonance microarrays,* Anal Biochem **351**, 241 (2006).

[284] D. Lakayan, R. Haselberg, R. Gahoual, G. W. Somsen, and J. Kool, *Affinity profiling of monoclonal antibody and antibody-drug-conjugate preparations by coupled liquid chromatography-surface plasmon resonance biosensing,* Anal Bioanal Chem **410**, 7837 (2018).

[285] D. Chen, E. N. McCool, Z. Yang, X. Shen, R. A. Lubeckyj, T. Xu, Q. Wang, and L. Sun, *Recent advances (2019–2021) of capillary electrophoresis-mass spectrometry for multilevel proteomics,* Mass Spectrom Rev (2021).

SUMMARY

Therapeutic proteins have been successfully developed for advancing medical treatments. They are usually large molecules produced by host cells and have a high degree of complexity compared to synthetic small molecule-based therapeutics. The complexity is mainly attributed to the heterogenic nature of post-translational modifications (PTMs). Glycosylation is one of the main drivers of protein heterogeneity. Since each modification may potentially impact the safety and efficacy, analytical methods for the structural and functional characterization of protein-based therapeutics are highly demanded.

Chapter 1 introduces an overview of therapeutic proteins. Common PTMs of therapeutic proteins are described with a focus on glycosylation. The functional relevance of glycosylation is exemplified for Immunoglobulins (Igs), a major class of therapeutic proteins. Further, common approaches and challenges for the analytical characterization are introduced. Mass spectrometry (MS) is a key technique for the analysis of therapeutic proteins. Thus, the fundamental ionization techniques for the analysis of biomolecules, electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) are described. Finally, the scope of the thesis is outlined.

The (glyco)peptide-centered analysis of therapeutic proteins is important for the site-specific characterization of PTMs. Bottom-up approaches are generally very data-rich, particularly for glycoproteins, due to the high structural diversity of glycosylation. **Chapter 2** presents a data analysis pipeline for streamlined processing of glycopeptide information from reversed-phase liquid chromatography MS/MS. A tryptic digest of IgG and IgA, enriched from human plasma, was analyzed in this study. Several recent bioinformatics tools, namely Byonic, GlycopeptideGraphMS and LaCyTools, are integrated. In addition, drawbacks and limitations of current glycoproteomics approaches are described. This chapter provides a guide for the efficient analysis of glycoproteomics data with a focus on maximizing glycopeptide identification and increasing the reliability of relative quantification.

The sample preparation strategy is highly relevant for a successful and efficient analysis of glycoproteins by bottom-up approaches. One of the main challenges is to obtain peptides bearing only a single glycosylation site. The addition of artificial protease cleavage sites via cysteine aminoethylation is exploited for the glycoproteomic analysis of recombinant erythropoietin (EPO) in **Chapter 3**. This approach adds additional tryptic cleavage sites between two glycosylation sites which could previously only be assessed by less specific proteases. It is demonstrated that the developed strategy, relying on trypsin as a sole protease, significantly reduced the data complexity and allowed for a shorter analysis time. Moreover, the advantage of the new method for

multiple attribute monitoring is presented.

Glycoproteomic approaches are important for an in-depth site-specific characterization of glycoproteins with multiple glycosylation sites. However, the information on the combination of glycoforms is lost. Valuable information on the actual proteoform heterogeneity is commonly obtained by intact mass profiling. ESI MS-based methods are usually the first choice for intact mass analysis due to higher proteoform resolution and increased stability of glycosylation features (i.e. sialic acids) compared to MALDI MS. **Chapter 4** demonstrates the feasibility of MALDI Fourier transform cyclotron resonance (FT-ICR) MS for the intact analysis of the complex and highly sialylated glycoprotein EPO. It is shown that the instrumental superiority and an appropriate choice of the MALDI matrix allow characterizing intact proteoforms by MALDI FT-ICR MS with comparable performance to ESI MS-based methods. The developed method is fast and offers a straightforward data analysis of doubly charged ions omitting the need for any charge deconvolution. This makes the approach highly relevant for developing high-throughput applications for the intact mass analysis of glycoproteins.

Monoclonal antibodies (mAbs) show a diversity of glycoforms (i.e., pairings of conserved Fc glycans) which may differentially impact the safety and efficacy. The structure-function characterization of mAb glycoforms is commonly addressed by glycoengineering of specific features and subsequent analysis by binding assays to relevant immune receptors. This approach is time-consuming and limited by the purity and variety of glycoforms that can be achieved. Fc gamma receptor IIIA is an important receptor for assessing antibody-dependent cellular cytotoxicity. **Chapter 5** describes a novel method based on Fc gamma RIIIa affinity chromatography (AC) hyphenated to MS, which allows unpreceded insights into the structure-function relationship of mAb glycoforms. Individual glycoforms can be easily ranked by their affinity without the need for glycoengineering.

The high mAb glycoform resolution achieved by Fc gamma RIIIA AC-MS may be hampered for mAbs with additional heterogeneity (e.g. Fab glycosylation). Therefore, the analysis of Fc moieties by a middle-up approach is desirable to decrease the proteoform complexity for a more confident glycoform assessment. In **Chapter 6**, the Fc gamma RIIIA method is further developed by evaluating different hinge-cleavage proteases for middle-up Fc gamma RIIIA AC-MS. Kgp is found to be a suitable protease for retaining Fc gamma RIIIA affinity of Fc moieties. The advantages for reducing the ambiguities in glycoform assignments are demonstrated for the analysis of Fab glycosylated cetuximab.

Chapter 7 shows the development of a novel Fc gamma RIIIB AC-MS method. Fc gamma RIIIB plays a major role in neutrophil activation, which is an important mechanism of action of therapeutic mAbs. The use of an Fc gamma RIIIb AC column is reported for the first time. The applicability of the method to a panel of different classical and glycoengineered mAbs is demonstrated. Furthermore, this chapter

compares the glycosylation profiles of human and recombinant Fc gamma RIIIB. In addition, the IgG subclass specificity of Fc gamma RIIIb AC-MS is shown.

Finally, the analytical challenges for the characterization of therapeutic glycoproteins are discussed in **Chapter 8**. The first part describes perspectives towards higher data analysis efficiency and reliability of bottom-up and intact mass analysis. In addition, potential biopharmaceutical applications for cysteine aminoethylation and intact mass analysis by MALDI MS are discussed. In the second part, the role of AC-MS for supporting critical quality attribute assessment of mAbs is highlighted. Technological perspectives are described for enhancing the functional understanding of mAb proteoforms. Finally, technical aspects of AC-MS are discussed and compared to other binding assays for applications in the biopharmaceutical industry.

SAMENVATTING

Therapeutische eiwitten zijn met succes ontwikkeld voor de bevordering van medische behandelingen. Het zijn meestal grote moleculen die door gastheercellen worden geproduceerd en een hoge mate van complexiteit hebben in vergelijking met synthetische, kleine-moleculaire geneesmiddelen. De complexiteit wordt voornamelijk toegeschreven aan de heterogene aard van post-translationele modificaties (PTMs). Glycosylering is een van de voornaamste oorzaak van de heterogeniteit van eiwitten. Aangezien elke modificatie mogelijk de veiligheid en werkzaamheid van therapeutische eiwitten kan beïnvloeden is er een grote behoefte aan analytische methoden voor de structurele en functionele karakterisering van eiwit-gebaseerde therapeutica.

Hoofdstuk 1 introduceert een overzicht van therapeutische eiwitten. Veel voorkomende PTMs van therapeutische eiwitten worden beschreven met een focus op glycosylering. Verder wordt de functionele relevantie van glycosylering beschreven voor een belangrijke klasse van therapeutische eiwitten, namelijk Immunoglobulinen (Ig). Daarnaast beschrijft dit hoofdstuk veelgebruikte technieken en uitdagingen voor de analytische karakterisering van Ig. Massaspectrometrie (MS) speelt een sleutelrol in de analyse van therapeutische eiwitten. Daarom geeft dit hoofdstuk uitleg over de fundamentele ionisatietechnieken voor de analyse van biomoleculen, namelijk *electrospray ionization* (ESI) en *matrix-assisted laser desorption ionization* (MALDI). Tenslotte wordt in dit hoofdstuk de reikwijdte van het proefschrift geschetst.

De analyse van therapeutische eiwitten op glycopeptide niveau is essentieel voor de locatie-specifieke karakterisering van PTMs. Bottom-up benaderingen zijn over het algemeen zeer datarijk voor glycoproteïnen, vanwege de hoge structurele diversiteit van glycosylering. **Hoofdstuk 2** presenteert een data-analyse pijplijn voor de gestroomlijnde verwerking van glycopeptide data verkregen door reversed-phase vloeistofchromatografie gekoppeld aan tandem MS (RPLC-MS/MS). Dit hoofdstuk beschrijft de analyse van een tryptische digest van IgG en IgA, verrijkt uit menselijk plasma. Verschillende recente bioinformatica tools, zoals Byonic, GlycopeptideGraphMS en LaCyTools, zijn geïntegreerd. Daarnaast beschrijft dit hoofdstuk de nadelen en beperkingen van de huidige glycoproteomics benaderingen. Dit hoofdstuk dient als leidraad voor de efficiënte analyse van glycoproteomics data. Tevens is dit hoofdstuk gefocust op het maximaliseren van het aantal geïdentificeerde glycopeptiden en het verhogen van de betrouwbaarheid van relatieve kwantificering.

De monstervoorbewerking is een zeer essentiële stap voor een succesvolle en efficiënte analyse van glycoproteïnen door bottom-up benaderingen. Een van de belangrijkste uitdagingen is het verkrijgen van peptiden met slechts één enkele glycosyleringslocatie. In **hoofdstuk 3** wordt een monstervoorbewerkingsstrategie ontwikkeld waarbij

kunstmatige protease splitingslocaties worden toegevoegd via cysteïne aminoethylering voor de glycoproteomische analyse van recombinant erytropoëtine (EPO). Deze strategie resulteert in extra tryptische splitsingslocaties tussen twee glycosyleringslocaties die voorheen alleen door minder specifieke proteasen werden verkregen. In dit hoofdstuk wordt aangetoond dat de ontwikkelde monstervoorbewerking, gebaseerd op trypsine als enige protease, de complexiteit van de data aanzienlijk vermindert en resulteert in een kortere analysetijd. Tot slot beschrijft dit hoofdstuk het voordeel van de nieuwe monstervoorbewerking voor *multiple attribute monitoring*.

Glycoproteomische benaderingen zijn belangrijk voor de locatie-specifieke karakterisering van glycoproteïnen met meerdere glycosyleringslocaties. Echter gaat hierdoor de informatie over de combinatie van glycoformen verloren. Daarom wordt waardevolle informatie over de werkelijke heterogeniteit van de proteoform meestal verkregen door intacte massa profilering. ESI-MS-gebaseerde methoden zijn doorgaans de eerste keuze voor intacte massa analyse. Dit is vanwege de hogere proteoform resolutie en grotere stabiliteit van glycosylatie kenmerken (bijv. siaalzuren) in vergelijking met MALDI MS. In **hoofdstuk 4** wordt de haalbaarheid van MALDI Fourier transform cyclotron resonance (FT-ICR) MS voor de intacte analyse van de complexe en sterk gesialyseerde glycoproteine EPO aangetoond. De instrumentele superioriteit en geschikte MALDI matrix resulteert in de karakterisering van intacte proteoformen met MALDI FT-ICR MS met vergelijkbare prestaties als ESI-MS-gebaseerde methoden. De ontwikkelde methode is snel en biedt een eenvoudige data-analyse van dubbel geladen ionen, zonder de noodzaak elke lading te deconvoleren. Dit maakt deze methode zeer relevant voor het ontwikkelen van high-throughput toepassingen voor de intacte massa analyse van glycoproteïnen.

Monoklonale antilichamen (mAbs) vertonen een diversiteit aan glycoformen (d.w.z., paren van geconserveerde Fc glycanen) die een verschillende invloed kunnen hebben op de veiligheid en werkzaamheid van het geneesmiddel. De structuur-functie relatie van mAb glycoformen worden meestal bepaald door glycoengineering van specifieke kenmerken, gevolgd door binding assays aan relevante immuunreceptoren. Deze aanpak is tijdrovend, en beperkt de zuiverheid en de verscheidenheid van glycoformen die kunnen worden behaald. Fc gamma receptor IIIA is een belangrijke receptor voor de bepaling van *antibody-dependent cellular cytotoxicity*. **Hoofdstuk 5** beschrijft een nieuwe methode op basis van Fc gamma RIIIa affiniteitchromatografie (AC) gekoppeld aan MS, die ongekende inzichten geeft in de structuur-functie relatie van mAb glycoformen. Aan de hand van deze methode kunnen individuele glycoformen gemakkelijk worden gerangschikt op basis van hun affiniteit zonder de noodzaak voor glycoengineering.

De resolutie van een mAb glycoform in Fc gamma RIIIA AC-MS kan worden belemmerd voor mAbs met extra heterogeniteit (bijv. Fab glycosylatie). Daarom is de analyse van Fc samenstellingen door een middle-up benadering wenselijk, aangezien dit de proteoform complexiteit vermindert en leidt tot een meer betrouwbare glycoform

beoordeling. In **hoofdstuk 6** wordt de Fc gamma RIIIA methode verder ontwikkeld door het evalueren van verschillende hinge-cleavage proteasen voor middle-up Fc gamma RIIIA AC-MS. Kgp blijkt een geschikt protease te zijn voor het behouden van de Fc gamma RIIIA affiniteit van Fc moieties. Verder worden de voordelen voor het verminderen van ambiguïteiten in de glycoform bepaling gedemonstreerd voor de analyse van Fab geglycosyleerd cetuximab.

In **hoofdstuk 7** wordt de methode ontwikkeling van een nieuwe Fc gamma RIIIB AC-MS methode beschreven. Fc gamma RIIIB speelt een belangrijke rol in neutrofiele activatie, wat een belangrijk werkingsmechanisme is van therapeutische mAbs. Voor de allereerste keer wordt het gebruik van een Fc gamma RIIIb AC kolom gerapporteerd. In dit hoofdstuk wordt de toepasbaarheid van de methode aangetoond door het analyseren van verschillende klassieke en glycoengineerd mAbs. Verder vergelijkt dit hoofdstuk de glycosylatieprofielen van humaan en recombinant Fc gamma RIIIB. Tot slot wordt in dit hoofdstuk de IgG subklasse specificiteit van Fc gamma RIIIb AC-MS aangetoond.

Tenslotte volgt er een algemene discussie over de analytische uitdagingen voor de karakterisering van therapeutische glycoproteïnen in **hoofdstuk 8**. Het eerste deel van dit hoofdstuk beschrijft de perspectieven voor een hogere efficiëntie en betrouwbaarheid van de data-analyse voor bottom-up en intacte massa analyses. Daarnaast worden potentiële biofarmaceutische toepassingen voor cysteïne aminoethylering en intacte massa analyse door MALDI MS besproken. In het tweede deel van dit hoofdstuk wordt de rol van AC-MS voor de ondersteuning van de kritische kwaliteitsbeoordeling van mAbs toegelicht. Tevens worden de technologische perspectieven voor het verbeteren van het functionele begrip van mAb proteoformen besproken. Als laatste worden de technische aspecten van AC-MS besproken en vergeleken met andere bindingsassays voor toepassingen in de biofarmaceutische industrie.

ACKNOWLEDGEMENTS

(...) Nothing of significance was ever achieved by an individual acting alone. (. . .)

— John C. Maxwell, 2006

I am very grateful to have had the opportunity to pursue my PhD program at the CPM. First, I want to acknowledge the great support from my promotor, Manfred. Thank you so much for your PhD position offer(s) and your contribution to my personal development.

In addition, I would like to express my gratitude to my co-promotors Noortje and David. Both of you had a major impact on shaping my scientific mindset. I have learned a lot from you and I appreciate the great work environment we have had. You gave me the freedom to develop my ideas and guided me in writing manuscripts. Furthermore, I want to thank Elena and Simone, who had also great contributions to my scientific development.

Of course, every member of the CPM and especially the glycogroup contributed directly or indirectly to this thesis and made my PhD journey a great pleasure! The technicians (Agnes, Carolien, Lisa, Jan, Hans) have been crucial for my technical support! Also, I would like to thank my students Coen and Raashina for their great work. Besides the scientific aspects, I enjoyed the young and international group at the CPM. The borrels have been always special and fun! Besides acquiring skills at the lab bench, I could learn a lot from my co-workers, e.g. how to cook authentic Chinese dishes (special thanks to Di and Wei). And of course, I want to thank the people from my office, particularly Guusje, Sanne and Christoph, for creating a great work environment and organizing all the great after-work activities. In particular, I wanted to acknowledge Christoph, who was my companion from day one in my PhD journey. I was truly privileged to have the opportunity to go through the PhD with you. I really enjoyed all our social activities, scientific discussions and travels. You made this PhD experience exceptionally fun and greatly contributed to inspiring and motivating me.

Thank you Willem for being my paranymph when I will defend this work. And also for all our great activities in Katwijk, especially all the beach walks.

Marie-Curie fellowships involve a lot of traveling and courses. I had an amazing time with all the A4B fellows during our workshops exploring the EU! In particular, I would like to thank Richard and Max for their great supervision and support during my secondment at Ludger.

I would like to give my acknowledgments to all colleagues and collaborators who contributed to the work in this thesis. In particular, the great collaboration with Roche/Genentech (especially Dietmar, Markus, Alexander, Tilman) was one of the main drivers for the success of this work. I am very thankful for all opportunities and discussions. At this point, I want also to acknowledge my former supervisor Henning, who prepared me for the PhD and had a profound impact on my interest in the development of analytical methods for biopharmaceuticals. In addition, I would like to thank Mathew for his great support on GlycopeptideGraphMS as well as Peter for data analysis discussions and Gestur for expanding my immunological knowledge.

Zum Schluss möchte ich noch herzlichst meiner Familie danken. Ohne euch, Mama und Papa, wäre diese Thesis nicht möglich gewesen. Ich bin dankbar dafür, dass Ihr mich in allen Situationen und meinen Entscheidungen unterstützt. Auch dir, Steffi, vielen Dank. Du hast mich dazu inspiriert und motiviert eine Karriere in den Lebenswissenschaften anzustreben.

CURRICULUM VITÆ

Steffen Lippold was born on the 19th of April, 1992 in Goslar, Germany. He obtained his Abitur in 2011 from the Grosse Schule in Wolfenbüttel. Afterward, he started his Bachelor's studies in Biology at the Technische Universität Braunschweig. During this time, he participated in the International Genetically Engineered Machine student competition. After his graduation in 2014, he enrolled in a Master's program in Biochemistry/Chemical Biology at the Technische Universität Braunschweig. His study focus was the development of drugs. In addition to his regular curriculum, he performed two industrial internships at Roche Basel and Kaiseraugst, Switzerland. During his first internship, he worked on polysorbate analysis within a technical development department, which resulted in his first scientific publication. In his second internship, Steffen worked in a quality control department for large molecules. There, he gained experience in the analytical methods used for the release and stability testing of antibody-drug conjugates. Both internships profoundly contributed to his interest in analytical methods for the characterization of biopharmaceuticals. After his graduation with honors, Steffen started a PhD program in 2018, under the supervision of Dr. Noortje de Haan, Dr. David Falck and Prof. Dr. Manfred Wuhrer, at the Center for Proteomics and Metabolomics, Leiden University Medical Center, The Netherlands. His PhD program was embedded in the Horizon 2020 Marie Sklodowska-Curie Action innovative training network Analytics for Biologics. In this project, Steffen focused on the development of novel mass spectrometry-based methods for the analysis of therapeutic glycoproteins. He gained comprehensive knowledge in the multi-level characterization of protein glycosylation, which led to six first-author manuscripts and this thesis. Next, Steffen will start a postdoctoral position at Genentech in California, where he will further explore the potential of affinity chromatography-mass spectrometry for the critical quality attribute assessment of biopharmaceuticals.

LIST OF PUBLICATIONS

1. **Impact of mono- and poly-ester fractions on polysorbate quantitation using mixed-mode HPLC-CAD/ELSD and the fluorescence micelle assay**
   S. Lippold, S.H.S. Koshari, R. Kopf, R. Schuller, T. Buckel, I.E. Zarraga, H. Koehn
   Journal of Pharmaceutical and Biomedical Analysis (132), 23-34, 2017.

2. **Glycoform-resolved FcγRIIIa affinity chromatography–mass spectrometry**
   S. Lippold, S. Nicolardi, E. Domongiuez-Vega, A.K. Heidenreich, G. Vidarsson, D. Reusch, M. Haberger, M. Wuhrer, D. Falck
   mAbs 7 (11), 1191-1196, 2019.

3. **Proteoform-Resolved FcγRIIIa Binding Assay for Fab Glycosylated Monoclonal Antibodies Achieved by Affinity Chromatography Mass Spectrometry of Fc Moieties**
   S. Lippold, S. Nicolardi, M. Wuhrer, D. Falck
   Frontiers in Chemistry 7:698, 2019.

4. **Cysteine Aminoethylation Enables the Site-Specific Glycosylation Analysis of Recombinant Human Erythropoietin using Trypsin**
   S. Lippold, A. Buettner, M.S.F. Choo, M. Hook, C.J. de Jong, T. Nguyen-Khuong, M. Haberger, D. Reusch, M. Wuhrer, N. de Haan
   Analytical Chemistry 92 (14), 9476-9481, 2020.

5. **Semiautomated glycoproteomics data analysis workflow for maximized glycopeptide identification and reliable quantification**
   S. Lippold, A.H. de Ru, J. Nouta, P.A. van Veelen, M. Palmblad, M. Wuhrer, N. de Haan
   Beilstein Journal of Organic Chemistry 16, 3038-3051, 2020.

6. **Glycoform Analysis of Intact Erythropoietin by MALDI FT-ICR Mass Spectrometry**
   S. Lippold, R. Thavarajah, D. Reusch, M. Wuhrer, S. Nicolardi
   Analytica Chimica Acta, 1185, 339084, 2021

7. **Fc gamma receptor IIIb binding of individual antibody proteoforms resolved by affinity chromatography–mass spectrometry**
   S. Lippold, A. Knaupp, A.H. de Ru, R.T. N. Tjokrodirijo, P.A. van Veelen, E. van Puijenbroek, S.W. de Taeye, D. Reusch, G. Vidarsson, M. Wuhrer, T. Schlothauer, D. Falck
   mAbs *in press*